EXPLOITING THE POSSIBILITIES OF SIMULATORS FOR DRIVER TRAINING

Proefschrift

Ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft;
op gezag van de Rector Magnificus prof. ir. K.Ch.A.M. Luyben
voorzitter van het College voor Promoties
in het openbaar te verdedigen op dinsdag 26 november 2013 om 10.00 uur
door Stefanus DE GROOT
Ingenieur luchtvaart en ruimtevaart
geboren te Haarlem

Dit proefschrift is goedgekeurd door de promotoren:
Prof. dr. ir. P.A. Wieringa
Prof. dr. ir. M. Mulder

Copromotor: Dr. ir. J.C.F. de Winter

Samenstelling promotiecommissie:

Rector Magnificus, voorzitter
Prof. dr. ir. P.A. Wieringa, Technische Universiteit Delft, promotor
Prof. dr. ir. M. Mulder, Technische Universiteit Delft, promotor
Dr. ir. J.C.F. de Winter, Technische universiteit Delft, copromotor
Prof. dr. ir. S.P. Hoogendoorn, Technische Universiteit Delft
Prof. dr. P.A. Hancock, Universiteit van Centraal Florida
Dr. J.P. Davis, Dragonfly Technology Ltd.
Dr. W.P. Vlakveld, Stichting Wetenschappelijk Onderzoek Verkeersveiligheid

Prof. dr. ir. J.W.C. van Lint, Technische Universiteit Delft, reservelid

Een groot deel van dit onderzoek is gesubsidieerd door het Nederlands Ministerie van Economische Zaken, onder het innovatiegerichte onderzoeksprogramma Mens-Machine Interactie, IOP MMI. Titel project: Virtual Assistant.

Ir. J. Kuipers (Green Dino BV, Nederland) heeft in belangrijke mate bijgedragen aan de initiatie en ondersteuning van dit onderzoeksproject.

Contents

Samenvatting

Het benutten van de mogelijkheden van simulatoren voor rijtraining

Trainen met een simulator biedt voordelen ten opzichte van trainen in een niet gesimuleerde omgeving. Over het algemeen is het goedkoper, veiliger, is er controle over de trainingsomgeving, en kan er gemakkelijk (zonder dure sensoren) data over de toestand van het voertuig en de bestuurder worden verzameld en opgeslagen. Door deze voordelen hebben simulators potentieel om effectieve training te bewerkstelligen.

Dit proefschrift bestaat uit drie delen, gericht op het realiseren van kosteneffectieve rijtraining en prestatiebeoordeling met een simulator. Het eerste deel richt zich op de prestaties van bestuurders in niet-bewegende rijsimulatoren. Het tweede deel richt zich op het leren autorijden in een rijsimulator, en het derde deel richt zich op de evaluatie van een nieuw ontwikkelde simulator voor het trainen en beoordelen van autocoureurs.

Het eerste deel van dit proefschrift onderzoekt of rijprestaties van bestuurders in vaste-basis rijsimulatoren realistischer gemaakt kunnen worden. In Hoofdstuk 1 wordt onderzocht welke prestatiematen het meest geschikt zijn voor het beoordelen van de rijprestaties tijdens het remmen en tot stilstand komen bij een stopteken. Remmen en tot stilstand komen is een taak die relatief veel gebruikt wordt in onderzoek naar rijgedrag en het beoordelen van de prestaties van bestuurders, en er worden veel verschillende maten gebruikt om deze rijtaak te analyseren. We beginnen met een analyse van de maten die in de literatuur gerapporteerd worden, met als doel te bepalen welke maten in het vervolg van dit proefschrift gebruikt kunnen worden. Er is eerst een computersimulatie uitgevoerd, waarna empirisch verkregen resultaten gebruikt worden om een aantal maten te analyseren. De belangrijkste conclusies zijn dat betrouwbare en valide maten voor de rem manoeuvre zijn: snelheid en afstand tot de stopstreep op het moment dat begonnen wordt met remmen, de stoppositie ten opzichte van het stopteken, en een maat die aangeeft of er met een constant acceleratie geremd is of met schokken.

In Hoofdstuk 2 worden acht acceleratie- en snelheidsfeedback systemen getest: een aanspannende riem, een vibrerend stuurwiel, een bewegende stoel, piepend bandengeluid, auditieve beeps, een trilmatje in twee verschillende configuraties, en een drukstoel. Voor de vijf systemen die longitudinale versnellingen terugkoppelen aan de bestuurder zijn de aanbevolen maten van Hoofdstuk 1 gebruikt. De conclusies zijn dat de voertuigacceleraties naar de bestuurder teruggekoppeld kunnen worden zonder gebruik te maken van een bewegingsplatform. Het systeem waarbij de riem aanspant als gevolg van longitudinale versnellingen zorgt voor een grote verbetering in het realisme van de rijprestaties.

In Hoofdstuk 3 is de modaliteit waarmee instructies in de simulator worden aangeboden onderzocht. Over het algemeen worden instructies in rijsimulatoren via verbaal aangeboden. De auditieve modaliteit is voor de hand liggend, omdat autorijden wordt gezien als een visuele taak waar de verbale instructies minimaal mee interfereert. Echter, beginnende leerlingen krijgen veel verbale instructies in een korte tijd. Het is daarom interessant om te onderzoeken wat het effect is van het aanbieden van route-instructies in een andere modaliteit. Het experiment in de rijsimulator liet zien dat zowel visuele als visueel-auditieve route-instructies

resulteerden in minder foute afslagen dan auditieve route-instructies. De visueel-auditieve instructies resulteren bovendien in een snellere reactietijd voor het aanzetten van de richtingaanwijzer. De visueel-auditieve instructies werden het meest duidelijk gevonden door mensen die sneller reden, of mensen die zelf rapporteerden dat ze niet goed konden rijden. De meeste mensen hadden een voorkeur voor visuele route-instructies. Dit experiment laat zien dat ondanks dat visuele instructies volgens de 'multiple resource theory' (Wickens, 1999) interfereren met de visuele rijtaak, het toch resulteerde in verbeterde prestaties.

Het tweede deel van dit proefschrift richt zich op het leren autorijden in een rijsimulator. Eerst worden de didactische eigenschappen van vier in Nederland verkrijgbare rijsimulatoren voor commerciële rijtrainingen geanalyseerd. Deze analyse is gedaan aan de hand van de 'first principles of instruction' (Merrill, 2002a), ofwel 'de grondbeginselen van instructie'. Er is ook gekeken naar de mogelijkheden van rijsimulatoren om een lesprogramma aan de grondbeginselen te laten voldoen. Geconcludeerd wordt dat geen enkele van de huidige commercieel verkrijgbare rijsimulatoren aan alle grondbeginselen voldoet, maar dat het in principe wel mogelijk is om een lesprogramma aan alle grondbeginselen te laten voldoen. Geconcludeerd wordt verder dat de visualisatie-, demonstratie-, en prestatiebeoordeling-mogelijkheden die geboden worden door moderne rijsimulatoren benut kunnen worden om de training te verbeteren. De objectieve meetgegevens uit de simulator kunnen gebruikt worden voor accurate en precieze feedback op de prestaties, iets dat praktisch niet mogelijk is in echte auto's, maar wat wel belangrijk is om efficiënte training te verzorgen. Er wordt aanbevolen om met gerichte experimenten het ontwerp van de training stapsgewijs te verbeteren.

In de drie volgende experimenten is onderzocht of er, door gebruik te maken van de potentiële voordelen die simulators bieden, rijvaardigheden bijgebracht kunnen worden in de rijsimulator. Hiervoor worden enkele psychologische fenomenen aangaande 'augmented feedback' (feedback van extra of versterkte informatie over de taak) bestudeerd en gebruikt om een nieuwe leeromgeving te creëren. In het eerste experiment worden geautomatiseerde stoelvibraties als functie van de laterale rijbaanpositie gebruikt met het doel onervaren bestuurders te leren om nauwkeurig in het midden van de rijbaan te rijden. Er zijn vier experimentele trainingscondities vergeleken: (a) on-target: stoel vibreert als het midden van de auto binnen 0.5 m van het midden van de rijbaan is, (b) off-target: stoel vibreert als het midden van de auto meer dan 0.5 m van het midden van de rijbaan verwijderd is, (c) controle: geen extra tactiele feedback, en (d) realistisch: stoel vibreert als functie van het motortoerental. Er zijn vijf korte ritten gereden: drie trainingsritten en twee retentieritten, waarvan de eerste retentierit direct na de trainingsritten, en de tweede retentierit een dag later. Tijdens de twee retentieritten werd gereden met de realistische stoelvibraties. De twee groepen met geaugmenteerde feedback, on-target en off-target, presteerden tijdens de training beter dan de twee groepen die geen geaugmenteerde feedback hadden. Tijdens de retentie-sessies was dit effect echter verdwenen. Tijdens de laatste training en beide retentiesessies presteerde de off-target groep beter in het rijbaan houden dan de on-target groep. De conclusie van dit experiment is dat off-target feedback effectiever is voor training dan on-target feedback.

Tijdens de twee volgende experimenten is de moeilijkheidsgraad van de simulatortraining aangepast door de grip van de banden tijdens de training te beïnvloeden. Het eerste experiment is uitgevoerd met een normale auto, het tweede met een raceauto. Een literatuuronderzoek laat zien dat het moeilijker maken van de taak tijdens de training het leren van die taak op langere termijn kan verbeteren. Net als tijdens het vorige experiment was de taak van de bestuurders om zo goed mogelijk rijbaan te houden. Er werden drie groepen met elkaar vergeleken: lage grip (LG), normale grip (NG), en hoge grip (HG). Na de vier trainingssessies volgden twee retentiesessies waarin met normale grip gereden werd. De LG groep reed tijdens de training en retentiesessies langzamer dan NG, had tijdens de training veel problemen om de auto op de weg te houden, maar tijdens de retentiesessies juist niet. In de overgang van training naar retentie nam de werkbelasting, gemeten met een auditieve reactietaak, van LG af en van HG juist toe. Dit experiment geeft verder aan dat het mogelijk is het gerapporteerde zelfvertrouwen van leerlingen te beïnvloeden, en dit heeft mogelijk implicaties voor de manier waarop rijles gegeven wordt.

In het tweede experiment waarbij de grip tijdens de training gemanipuleerd is onderzoeken we niet reguliere autobestuurders, maar autocoureurs. Nu is het doel niet om langzamer te rijden, maar juist sneller. Drie groepen onervaren coureurs zijn getraind en getest op hetzelfde eenvoudige racecircuit: lage grip (LG), normale grip (NG), en hoge grip (HG). Net als bij het vorige experiment reed LG langzamer dan de andere groepen in zowel de trainingen als de eerste retentiesessie. De tweede retentiesessie werd met een andere auto gereden dan de trainingen en eerste retentiesessie (Formule 1 in plaats van Formule 3 auto), en in deze sessie werden geen significante verschillen in rondetijd gemeten. LG had echter wel een lagere zelf-gerapporteerd frustratie en had het idee sneller te rijden dan de gemiddelde proefpersoon uit zijn/haar groep.

Het derde deel van dit proefschrift richt zich op de validiteit en controleerbaarheid van een race-simulator. Er is een bescheiden validatiestudie van de race-simulator gedaan door de snelste rondetijden van 13 coureurs tijdens trainingssessies in de simulator zijn vergeleken met de snelste rondetijden van dezelfde coureurs tijdens trainingen in de echte wereld. Er is een significant verband tussen de rondetijden in de simulator en de echte wereld gevonden, wat suggereert dat de simulator een zekere voorspellende waarde heeft voor prestaties op het echte circuit.

Een (race)simulator kan gebruikt worden voor goed gecontroleerde tests die in de werkelijkheid moeilijk uit te voeren zijn. We hebben bij verschillende raceauto's grote verschillen in de versterkingsfactor (gain) en stijfheid van het rempedaal gevonden. Het vermoeden bestaat dat er een optimale stijfheid en gain bestaat voor autocoureurs, maar dit is lastig in de werkelijkheid uit te vinden: De te verwachten prestatieverschillen zijn klein, de tijd om het remsysteem om te bouwen lang, en de omgevingsfactoren zoals de grip van het circuit en de banden veranderen continu. In twee onafhankelijke experimenten is het effect van de stijfheid van het rempedaal op de rondetijd onderzocht. We verwachtten dat een slap rempedaal voor snellere sectortijden zou zorgen als een sector een lange remzone bevatte, en dat een stijf pedaal zou resulteren in meer hoog frequente rem-inputs van de coureur. Uit deze experimenten blijkt dat autocoureurs adaptief zijn en kunnen omgaan met zeer uiteenlopende remstijfheden, en dat een stijf rempedaal voor snellere controle-inputs zorgt. De rijsimulator bleek opnieuw een bruikbaar hulpmiddel om experimenten aan de mens-machine interface uit te voeren die in de echte wereld moeilijk te realiseren zijn.

Om in meer detail te achterhalen welke eigenschappen van het rempedaal belangrijk zijn voor het controleren van de remkracht door autocoureurs is de versterkingsfactor van het rempedaal nader onderzocht. Bij dit experiment wordt niet gereden over een virtueel circuit, maar wordt door deelnemers een eendimensionale controletaak uitgevoerd. Het rempedaal is isometrisch geconfigureerd, wat inhoudt dat er nauwelijks verplaatsing optreedt en dat de gemeten remkracht bepalend is voor de uitvoer. Op deze manier zijn vier versterkingsfactoren vergeleken, variërend van zeer laag tot zeer hoog, in twee verschillende controletaken. Wat betreft nauwkeurigheid van taakprestatie waren de effecten van de versterkingsfactor klein, maar er werden sterke effecten gevonden wat betreft de opbouwsnelheid, overshoot, variabiliteit binnen personen, en zelf-gerapporteerde fysieke werklast. De resultaten bevestigen dat het vinden van de optimale versterkingsfactor een afweging is tussen stabiliteit en werklast.

Summary

Exploiting the possibilities of simulators for driver training

Training in a simulator offers potential advantages compared to training in a non-simulated environment. Generally it is cheaper, safer, there is more control over the environment, and data collection is less complicated. These potential advantages give simulators the possibility to offer effective training.

This thesis is divided into three parts, aimed at realizing cost-effective driver training and driver assessment using simulators. The first part focuses on driver performance in fixed-base simulators, the second part focuses on learning to drive in a simulator and the third part evaluates a newly developed simulator for the training and assessment of racecar drivers.

Valid and reliable performance measures are required to analyze driver performance. Chapter 1 evaluates a large amount of measures for the task of braking and then stopping at a stop-sign, which is a common task for research into driver performance. A computer simulation was executed and also empirical data was used to study the performance measures. The main conclusions were that reliable and valid measures for the braking task are: speed and distance to the stop-sign at the start of braking, the stopping position with respect to the stop-sign, and a measure which indicates whether or not the deceleration was constant while the vehicle was slowed down.

Chapter 2 tests eight low-cost non-vestibular acceleration and speed feedback systems: a tensioning seatbelt, a vibrating steering wheel, a motion seat, screeching tire sound, auditory beeps, a vibrating seat-pan in two configurations, and a pressure seat. For five systems, which provide longitudinal acceleration feedback, the measures of Chapter 1 were used to analyze the effect of the feedback systems on driver performance during the braking task. Chapter 2 concludes that vehicle acceleration cues can be fed back to the driver without a motion platform. The system which made the largest gain in making driver performance more realistic was the tensioning seatbelt system.

Chapter 3 investigates the modality with which instructions are presented in the simulator. Generally, instructions in simulators are presented verbally. The auditory modality is a logical choice because car driving can be seen as a predominantly visual task. However, beginner drivers receive a lot of verbal instructions in a limited amount of time, and therefore it is interesting to investigate the effects of presenting the route instructions in a different modality. The experiment in the driving simulator showed that both visual and visual-auditory route-instructions resulted in less turning errors than the auditory route-instructions. The visual-auditory instructions also reduced indicator reaction times. The visual-auditory instructions were preferred by people who drove faster, and people who had low self-reported driving skill. Most people preferred the visual instructions over the auditory instructions. This experiment showed that even though the visual instructions interfere with the predominantly visual driving task according to the 'multiple research theory' (Wickens, 1999), they did result in better driving performance.

The second part of this thesis focuses on learning to drive in a simulator. First, the didactical properties of four commercially available driving simulators are analyzed. A survey shows that the intelligent tutoring systems of current driver

training simulators are mostly imitating the human instructor and that the "first principles of instruction" (Merrill, 2002a) are not implemented to their full potential. Hence, there is ample room for improvement of the didactical properties by fully exploiting the many visualization, demonstration, and performance-assessment opportunities provided by modern driving simulators. Furthermore, objective performance ratings of students can be used to provide accurate and consistent feedback-on-performance, something that is not possible in real cars, but which is often essential for effective skills training. It is recommended to use empirical experimentations to improve the instructional design of simulator-based driver training for specific learning outcomes and validate the use of the first principles of instruction to facilitate learning.

The following three experiments investigate whether potential advantages which are offered by simulators can be used to teach driving skills to learner drivers. Some psychological principles concerning augmented feedback are studied and used to create a new learning environment. In the first experiment, seat vibrations which reacted to the lateral position in the lane were used to teach inexperienced drivers to drive in the middle of the right lane. There were four experimental groups: (a) on-target, receiving seat vibrations when the center of the car was within 0.5 m of the lane center; (b) off-target, receiving seat vibrations when the center of the car was more than 0.5 m away from the lane center; (c) control, receiving no vibrations; and (d) realistic, receiving seat vibrations depending on engine speed. During retention, all groups were provided with the realistic vibrations. Every participant drove five 8 minute sessions: three training sessions, one retention test directly after practice, and one retention test the following day. During practice, on-target and off-target groups had better lane-keeping performance than the nonaugmented groups, but this difference diminished in the retention phase. Furthermore, during late practice and retention, the off-target group outperformed the on-target group. The conclusion of this experiment is that off-target feedback is superior to on-target feedback for learning the lane-keeping task.

During the following two experiments, the difficulty of the training was varied by changing the friction coefficient of the tire on the road. The first experiment deals with a normal road-car, while the second experiment deals with a racing car.

Previous research in motor learning has shown that degrading the task conditions during practice can enhance long-term retention performance. Just like in the previous experiment, the driving task was keeping the car in the center of the right lane. The inexperienced drivers were divided into three groups: low grip (LG), normal grip (NG), and high grip (HG). All groups drove six sessions: four practice sessions, an immediate retention session, and a delayed retention session the following day. The two retention sessions were driven with normal-grip tires. The results show that LG drove with lower speed than NG during practice and retention. Transferring from the last practice session to the immediate retention session, LG's workload decreased, as measured with a secondary task, whereas HG's workload increased. This experiment also showed that it is possible to influence self-reported confidence level, which may have potential implications for the way drivers are trained.

In the second experiment in which the tire-road friction coefficient is varied during training, we are not investigating normal car driving, but racecar driving. Now the goal is not to make people drive slower, but faster instead. Three groups of inexperienced racecar drivers were trained and tested on the same simple racetrack: low grip (LG), normal grip (NG), and high grip (HG). Just like in the previous experiment, LG drove slower than the other groups during training and the first retention session. The second retention session was driven in a different car than the training and the first retention session (Formula 1 car instead of a Formula 3 car), and in this session no differences in lap time were found between the groups. LG reported a higher confidence and lower frustration than NG and HG after each of the two retention sessions. In conclusion, practicing with low grip, as compared to practicing with normal or high grip, resulted in increased confidence but slower lap times.

The third part of this thesis investigates the validity and controllability of a racing simulator. A modest validation study was performed by comparing the fastest lap times of 13 racing drivers during training sessions in the simulator to the fastest lap times these same drivers did on the same track in the real world. A correlation between the lap times was found, which indicates that the simulator has some degree of predictive value for performance in the real world.

A (racing) simulator can be used for controlled experiments which are difficult to perform in reality. In different racecars, we have found large differences in gain and stiffness of the brake pedal. We assume that there exists an optimal stiffness and gain of the brake pedal for racecar drivers, but this is hard to investigate in reality. The expected performance differences are small, the time it takes to adapt the brake system is lengthy, and the environmental factors, such as grip of the tires and track, vary all the time. In two independent experiments the effect of the brake pedal stiffness on lap times is investigated. The expectations were that a softer brake pedal would be better in long brake zones, and that a stiff pedal would result in faster control inputs by the driver. The conclusions of the two experiments are that racing car drivers can deal with a large range of brake pedal stiffness, that a stiff pedal results in faster control inputs, and that the simulator is a useful tool for experiments concerning the human-machine interface which are difficult to perform in reality.

To get a more detailed idea about which properties of the brake pedal are important for brake force control of racecar drivers, the gain of the brake pedal is investigated further in Chapter 9. During the last experiment participants did not drive on a virtual track, but performed a one dimensional control task. The test setup was a formula racing car cockpit fitted with an isometric brake pedal, which means that the pedal does not deflect under load and the pedal force determines the output. Four control-display gains, varying from very low to very high, were compared with two target functions; a step function and a multisine function. The control-display gain had only minor effects on root mean-squared error between output value and target value but it had large effects on build-up speed, overshoot, within-participants variability, and self-reported physical load. The results confirm the hypothesis that choosing an optimum gain involves balancing stability against physical effort.

Preface

Human-in-the-loop simulators are widely used for training and assessment of operators such as pilots, surgeons, industrial plant operators, and car drivers. For airline pilots, simulators have become a compulsory element in the curriculum. The widespread use of simulators for pilot training is not surprising when comparing the costs of running a flight simulator with the costs of running a jet airliner, or when comparing the risks involved when a dangerous maneuver or situation is practiced. The cost difference between these two modes of training (simulator vs. reality) is large and growing, as simulators are becoming less expensive (reductions in the price of computers and displays) while aircraft cost are increasing (price of aircraft and kerosene).

For car driving the cost difference between training in a real car and in a simulator is considerably smaller as compared to flying. Furthermore, simulation of ground vehicles provides more stringent challenges than airborne vehicle simulation regarding simulator fidelity and virtual environments. This means that for a long time it has been more difficult to obtain cost-effective simulator-based driver training than pilot training. Because of the ever-decreasing cost of simulators, however, cost-effective simulator-based driver training is becoming increasingly viable. At this moment over 100 simulators are used for commercial driver training in The Netherlands. These simulators are operated by driving schools and used in various ways to complement the standard driver training curriculum.

Training in a simulator offers potential benefits compared to training in a real vehicle: a) driver safety, and thus the opportunity to learn from errors, b) the state of the car is fully known, offering easy and objective data collection, c) feedback and instructions can easily be presented in multiple modalities, d) control over the training conditions, and e) cost savings per time unit, especially when the human instructor can be replaced with a virtual instructor.

This thesis investigates the possible increase in effectiveness of simulator-based driver training, by exploiting the above-mentioned potential advantages. It uses several paradigms from psychology to develop new methods and feedback systems for car driver training. Experiments were carried out to measure the effect of these methods and feedback systems on driving performance and training effectiveness.

Learning comprises more than maximizing task performance during training. According to Schmidt and Bjork (1992), training for a real-world task like car driving should aim at the following two aspects: a) the level of performance in the long term, and b) the capability to transfer training to related tasks and altered contexts. During real-world driving lessons, however, the instructor has to make sure that the student driver does not make disastrous errors and tries to improve the driver's performance as quickly as possible (Groeger & Banks, 2007). For simple motoric tasks, experimental psychological research has repeatedly shown that learning can be improved by providing *less* feedback during training, by reducing the amount of feedback as a function of skill level, and by providing difficult and varied training conditions. These modifications result in worse task performance during training but in higher performance during retention tests. This thesis investigates whether these learning principles, which have been shown to apply to simple motor tasks, also apply to more complex driving tasks. That is, whether less feedback and more difficult training conditions during training sessions results in improved retention performance.

This thesis contains three parts. Part 1 investigates driving performance in fixed-base driving simulators. A number of experiments reported in the literature have shown that drivers behave more realistically in a motion-base simulator than in a fixed-base simulator. Because motion systems generally are too expensive for commercial driver training, we presented vehicular acceleration feedback without using a motion-base in a fixed-base simulator. Chapter 1 analyses performance measures which can be used to analyze driver braking performance, and Chapter 2 investigates eight low-cost motion cueing systems in elementary braking and cornering tasks.

Part 2 of this thesis focuses on the training of car drivers and starts with an investigation of the didactical state of the art of four commercially available driver training simulators. Chapters 5, 6, and 7 focus on learning performance in the driving simulator. These three experiments follow the same paradigm: different groups of drivers complete a number of training sessions and are then tested during retention sessions. One group serves as the control group, while the training of the other group(s) takes place under manipulated conditions. In two of the three experiments we tested the groups again on the day after the training, in a so-called delayed retention session. Chapter 5 investigates an augmented feedback system which presents seat vibrations based on lateral lane position during practice, a so-called bandwidth feedback system. Chapter 6 trained learner drivers under challenging low-grip conditions. Chapter 7 also manipulated tire grip during practice. Whereas Chapter 6 focuses on normal road car driving, Chapter 7 focuses on racecar driver training.

Part 3 of this thesis focuses on racecar driving, a long term passion of me. More specifically, Part 3 focuses on the performance of racecar drivers in fixed-base simulators. Race car driving is expensive and involves more safety risk than normal car driving. Therefore, the aforementioned cost and safety advantages of a driving simulator as compared to the real-world are considerably larger than for normal car driver training. Additionally, in order to cut down the cost of competition, practicing in racecars is often prohibited by the race series legislators, so drivers have very little opportunity to practice and refresh their skills in their real-world racecar. These arguments, in combination with the ever-decreasing cost of simulators, have fuelled an increase in the number of simulators that are used for commercial racecar driver training. For research purposes, the racecar driving task is highly interesting. Contrary to real-world driving tasks, the performance of racecar drivers is easily measured: the primary task is to minimize lap time. Lap time operationalizes all the complex driver's behavior and cognition into one simple measure that is always available both in reality and in the simulator.

Chapter 8 starts with a validation study, followed by a study of the effect of the brake pedal stiffness on lap times. Such an investigation is highly difficult to carry out in a real car, because of constantly varying environmental and vehicle parameters, but it *is* possible in a driving simulator. Chapter 9 reduces the task complexity of the braking task and studies brake pedal control performance as a function of control-display gain. The goal of this experiment is to find the gain with which the best control performance can be obtained.

Part 1

Driver performance in fixed-base driving simulators

Chapter 1. An analysis of braking measures

Abstract

Braking to a full stop at a prescribed target position is a driving maneuver regularly used in experiments to investigate driving behavior or to test vehicle acceleration feedback systems in simulators. Many different performance measures have been reported in the literature for analyzing braking. These may or may not be useful to analyze the stopping maneuver, because a number of potential problems exist: 1) the scores on a measure may be insufficiently reliable, 2) the measure may be invalid, or 3) the measure may be strongly correlated with other measure(s).

A simulation and an empirical study were conducted to analyze various measures. From the simulation study it is concluded that: 1) a measure based on the speed vs. time relationship can be used to measure deviations from a constant deceleration (R^2), 2) minimum time-to-collision (TTC_{min}) is sensitive to target position offsets, and 3) mean TTC-dot can capture braking behavior characteristics, but the required definition of a begin and end sample-point for its calculation is a disadvantage.

The empirical study calculated a set of measures using the data of 60 participants who drove ten stopping maneuvers in a driving simulator. It is concluded that reliable and valid measures for a stopping maneuver are provided by the speed and distance to the target position at braking onset, the stopping position with respect to the target, and the R^2 measure to measure deviations from a constant deceleration. Recommended additional measures are: the mean speed of the complete braking maneuver, stopping position consistency, maximum deceleration, and onset jerk.

De Groot, S., De Winter, J. C. F., Wieringa, P. A., & Mulder, M. (2009). An analysis of braking measures. *Proceedings of the Driving Simulation Conference Europe*, Monte Carlo, Monaco, 233–243.

1.1 Introduction

Braking to a full stop is a driving maneuver regularly used in experiments to investigate driver behavior or to test acceleration feedback and support systems. During a typical braking maneuver participants should decelerate the car from driving speed and stop close to a stopping target. Our aim is to improve the analysis of the braking maneuver by testing a large number of measures found in literature and then specifying the most valid and reliable ones.

To introduce the braking maneuver and the critical time-points, Figure 1 illustrates a representative braking maneuver. At time = t_0 the throttle is fully released, after which the brake is pressed at time = t_1. The brake pedal is depressed further, until the vehicle reaches its maximum deceleration during the maneuver at time = t_2. Finally, the vehicle comes to a complete stop at time = t_3.

Many different performance measures for analyzing braking are reported in the literature. Boer, Girshik, Yamamura, and Kuge (2000) compared the braking maneuver in a real car to the same maneuver performed in a simulator. They concluded that drivers braked later, harder, and in a multi-modal manner (multiple separate brake pedal applications) in the simulator rather than with a constant deceleration as was found in a real car. Following this comparison, Boer, Kuge, and Yamamura (2001) introduced a driver model for stopping behavior which could reproduce the multi-modal braking profiles. Differences between simulated and real driving behavior were also demonstrated by Jamson and Smith (2003). They fitted a second-order polynomial on the speed vs. distance data and quantified multi-modal braking using the R^2 measure.

Other braking experiments investigated the influence of motion platforms or low-cost motion cueing solutions on braking behavior (Brünger-Koch, Briest ,& Vollrath, 2006; De Winter, De Groot, Mulder, & Wieringa, 2007; De Winter, De Groot, Mulder, Wieringa, & Dankelman, 2008; Pinto, Cavallo, Ohlmann, Espié, & Rogé, 2004; Siegler, Reymond, Kemeny, & Berthoz, 2001). Both papers by De Winter et al. (2007, 2008) used the same measures as Siegler et al. (2001), with the exception of the distance to the target position at time = t_3. De Winter et al. (2007, 2008) adapted this measure to exclude inter-participant differences of the desired target position (stopping line or stopping sign) by using the standard deviation of the position error instead of the mean position error as used by Siegler et al. (2001). Brünger-Koch et al. (2006) calculated approach speed (t_0), total stopping distance (t_0–t_3), and the maximum deceleration (t_2), but also time-to-collision (T_c) at braking onset (t_1), total stopping time (t_0–t_3), and pedal transition time (t_0–t_1). They used time = t_0 as the start point for their calculations, where most other researchers have used time = t_1. Pinto et al. (2004) calculated TTC at the onset of braking, maximal deceleration, the instant of maximal deceleration ($t_2/(t_3-t_1)$), and braking smoothness measured by the number of inversions of the deceleration profile, total braking duration, and total stopping distance. All these papers report significant differences for the maximum deceleration during braking.

Research into the visual perception of speed and distance also focused on the braking maneuver. Researchers were inspired by the introduction of the tau (or visually obtained TTC) and tau-dot (or the time rate of change of TTC, TTC-dot) concepts of Lee (1976). This resulted in many experiments focusing on the braking maneuver to explore the concept of direct perception and time-based control (e.g., Flach, Smith, Stanard, & Dittman, 2003; Groeger, 2000; Van der Horst, 1990; Yilmaz & Warren, 1995).

Measures used to analyze the braking maneuver may or may not be useful, because a number of potential problems exist:

- The scores on a measure may be insufficiently reliable;
- A measure may be invalid, thereby not capturing the phenomenon of interest;
- When multiple measures are used, measures may be strongly intercorrelated and therefore not sufficiently diverse.

This paper evaluates performance measures obtained from literature and regards the braking maneuver by means of simulation and the analyses of experimental data. Our main goal is to investigate which measures are valuable for analyzing braking and to offer a better general understanding of the braking maneuver itself.

1.2 Simulation study

Before the measures were tested using experimental data, some complex measures required an additional simulation analysis to investigate the mathematical properties without the disturbances introduced by human operators. The measures evaluated during the simulation study were R^2, TTC_{min}, and mean TTC-dot (see Table 3 for further information concerning these measures).

Concerning R^2, Jamson and Smith (2003) intended to find a measure to quantify multi-modal braking, as Boer et al. (2000) defined it. They used a procedure to fit a second-order polynomial to the speed vs. distance graph, which resulted in what we define as the $R^2_{distance}$ measure. Because the speed vs. distance graph is not a second-order function with a constant deceleration, we propose another way to calculate this measure, in the time domain (R^2_{time}) and compare these two R^2 measures in the simulation study.

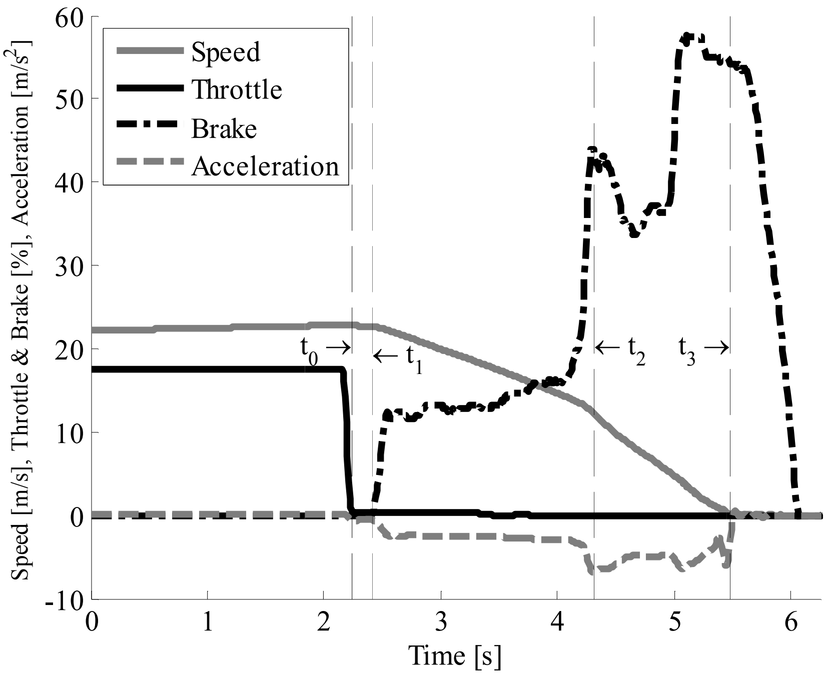

Figure 1. Definition of time-points during a single braking maneuver of a subject in a fixed-base driving simulator.

1.2.1 Method

The simulation modeled the movement of a car slowing down from a given initial speed with a prescribed deceleration profile. The simulation started at the moment the deceleration began (the brake onset). Eight braking maneuvers with different deceleration profiles were simulated to get a clear view of the impact of these diverse deceleration profiles on the calculated measures. An overview of the simulated cases is provided in Table 1. Figures 2a and 2b show the speed vs. time and the speed vs. distance graphs for the cases of Table 1. Cases 1 to 4 had constant deceleration profiles, from 30 and 80 km/h with different decelerations, whereas cases 5 to 8 had variable decelerations without (5, 6) or with (7, 8) modal braking.

1.2.3 Results

The results of $R^2_{distance}$ and R^2_{time} are presented in Table 2. For cases 1 to 4, R^2_{time} had a value of exactly 1.000, as expected with a constant deceleration, whereas $R^2_{distance}$ had a value slightly below one, showing the worse fit. The largest difference between these measures is found for case 8, where the driver is simulated to completely release the brake for a number of seconds before applying the brake again at low speed and close to the stop line. The speed vs. distance graph is shown together with a second-order fit in Figure 3a, showing the relatively good fit and thus high $R^2_{distance}$ score (0.981, see Table 2). Table 2 shows that R^2_{time} captures the multi-modal braking better than $R^2_{distance}$ for case 8 (0.683 vs. 0.981).

Table 1. Simulated cases.

Case	Speed (km/h)	Initial acceleration (m/s^2)	Acceleration rate of change (m/s^3)	Multi-modality (total number of brake releases)
1	30	−4	0	0
2	80	−4	0	0
3	80	−2	0	0
4	80	−8	0	0
5	80	−2	−2	0
6	80	−8	1	0
7	80	−4	0	5
8	80	−4	0	1

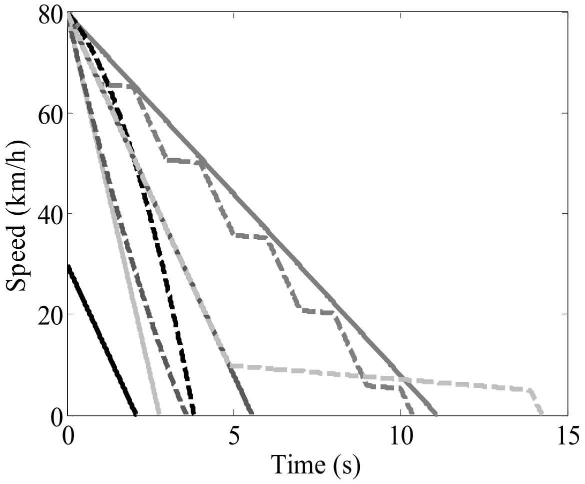

Figure 2a: Speed vs. time for the 8 simulated cases.

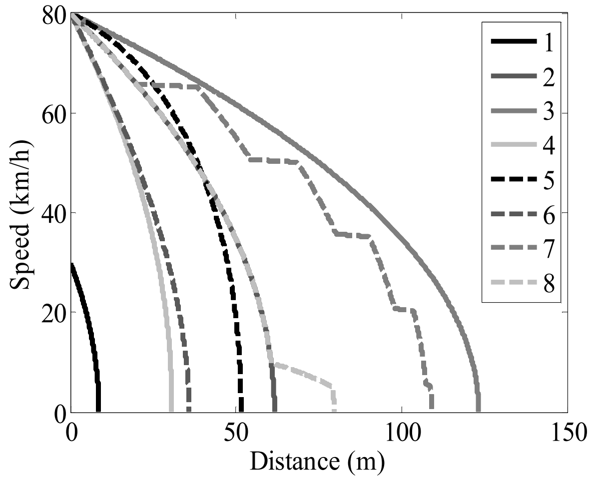

Figure 2b: Speed vs. distance for the 8 simulated cases.

Table 2 shows that the mean TTC-dot measure is able to distinguish between increased and decreased accelerations, as intended by Lee (1976) and also by Yilmaz and Warren (1995). The start- and end-point for the calculation of a mean TTC-dot (or the determination of a regression line such as Yilmaz & Warren) is critical for the measure's value. We used the TTC_{min} sample-point as end-point and brake onset as start-point of the calculation interval. We see that for case 1 (low speed) the nonlinear contribution of TTC-dot at the end of the calculation interval has a larger influence on the measure score than for case 3 (braking from a higher speed). Case 3 results in a mean value of −0.47 s, which is closest to the theoretical value of −0.50 s for a constant deceleration. If the stopping point is taken as endpoint, the mean TTC-dot value is infinite. Figure 3b illustrates the sensitivity of the TTC_{min} measure to a distance offset. It can be seen that the TTC_{min} measure is very sensitive to variations of the target stopping position, thereby reducing reliability and validity. When a participant interprets the stopping target differently (e.g., stop-sign instead of stop-line), or the person's head position instead of the front of the car is used to calculate the distance to the target, it is easy to obtain distance offsets of about 2 m.

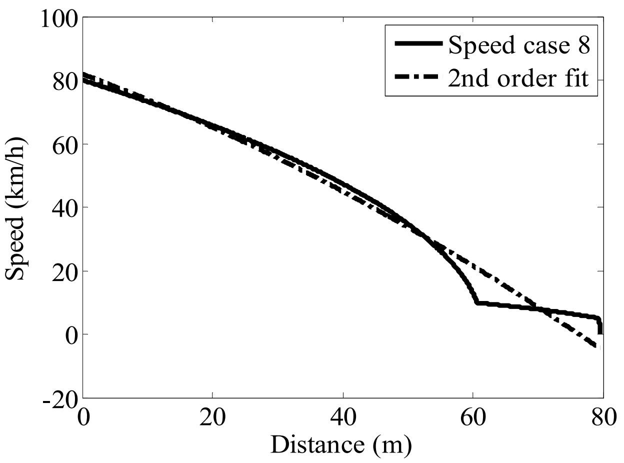

Figure 3a: Speed vs. distance profiles for case 2 with the second-order polynomial fit ($R^2_{distance}$).

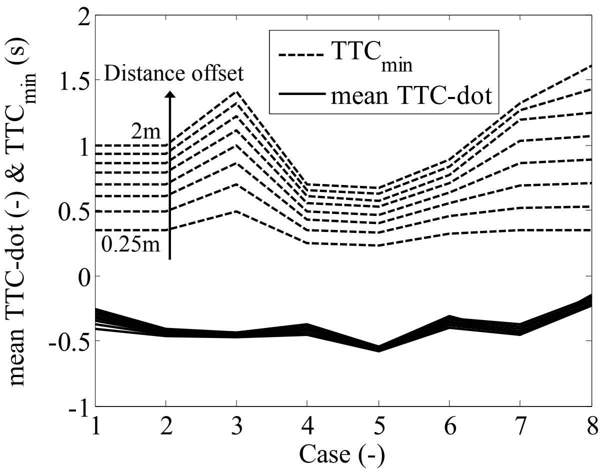

Figure 3b: TTC_{min} and mean TTC-dot for the 8 cases with varied target distance offset (range:0.25–2 m).

Table 2. Simulation results for the R^2, TTC_{min}, and mean TTC-dot measures

Case	$R^2_{distance}$	R^2_{time}	TTC_{min}	mean TTC-dot
1	0.993	1.000	0.50	−0.38
2	0.992	1.000	0.50	−0.45
3	0.993	1.000	0.70	−0.47
4	0.993	1.000	0.35	−0.44
5	0.992	0.972	0.32	−0.57
6	0.996	0.995	0.45	−0.38
7	0.979	0.992	0.52	−0.44
8	0.981	0.683	0.53	−0.22

Table 3. Measures used for braking analyses calculated for each participant.

Nr. ValidStops*,**	(-)	Number of successful braking maneuvers per session (De Winter et al., 2007)
Mean V	(m/s)	Mean speed of the braking maneuver (−175m < stopping position (t_3) < 30m) (De Winter et al., 2007)
V_{ini}	(m/s)	Speed at braking onset (t_1) (Siegler et al., 2001)
DTT_{ini}	(m)	Distance to target at braking onset (t_1) (Siegler et al., 2001)
DTT_{fin}	(m)	Distance to target at stopping position (t_3) (Siegler et al., 2001)
SD DTT_{fin}**	(m)	Standard deviation of distances to target at stopping position (t_3) (De Winter et al., 2007)
Max. dec.	(m/s^2)	Maximum deceleration during the maneuver (t_2) (Siegler et al., 2001)
Onset jerk	(m/s^3)	Mean jerk during the first part of the braking maneuver (t_1−(t_1-t_2)/2) (Siegler et al., 2001)
TTC_{ini}	(s)	Time-to-Collision (line-crossing) at the onset of braking (t_1) (Lee, 1976; Van der Horst, 1990)
TTC_{min}	(s)	Minimum Time-to-Collision (t_1−t_3) (Van der Horst, 1990)
Max. dec. pos.	(%)	Time of maximum deceleration with respect to the braking time (t_1−t_3) (Pinto et al., 2004)
Dec. inversions	(-)	Number of inversions of the deceleration profile (t_1−t_3) (Pinto et al., 2004)
SD dec.	(m/s^2)	Standard deviation of deceleration during braking (t_1−t_3)
TransferTime	(s)	Time it takes from throttle release to brake onset (t_1-t_0) (Brünger-Koch et al., 2006)
BrakeEvents	(-)	Number of brake applications after first braking onset (t_1−t_3)
ThrottleEvents	(-)	Number of throttle applications (t_1−t_3)
Max. Brake	(%)	Maximum brake position during braking (t_1−t_3)
Max. Throttle	(%)	Maximum throttle position during braking (t_1−t_3)
StopTime	(s)	Braking duration (t_3-t_1) (Pinto et al., 2004)
StopDistance	(m)	Stopping distance (t_1−t_3) (Pinto et al., 2004)
R^2_{time}	(-)	Squared correlation coefficient of the speed vs. time relationship (t_1−t_3) (adapted from Jamson & Smith, 2003)
mean TTC-dot	(-)	Mean TC-dot until the TTC_{min} (t_1−t_{TTCmin}) (adapted from Yilmaz & Warren, 1995)

*A successful stop was defined as a stop where the minimum speed is smaller than 1 km/hour.
**All other measures were calculated as the mean per subject of successful stops during one driving session.

1.3 Empirical study

1.3.1 Method

A set of braking measures (Table 3) was calculated for experimental data of which the results were partly published at previous DSC conferences (De Winter et al., 2007, 2008). These experiments compared the effects of various forms of motion cueing against a baseline condition without motion cueing during a brake maneuver. Only a participants' first driving session, and only when it was driven in the baseline condition, was included in the present paper. The total number of included sessions was 60, and each session comprised of 10 stopping maneuvers. The discriminant validity of the measures was determined using driving performance differences between experienced (defined as having a driving licence) and inexperienced drivers. Additionally, the first and last four stopping maneuvers were compared to check driver adaptation within a driving session. Differences between the measures were investigated using the *t* test and Cohen's *d* effect size measure, whereas a correlation matrix of all the measures was used to assess the redundancy of the measures. The split-half reliability of the measures was investigated by correlating the measures with themselves between the five even and five odd stop numbers.

1.3.2 Results

Tables 4 and 5 present the results of the empirical study. The results can be summarized as follows:

a) The experience comparison indicated a number of differences between experienced and inexperienced drivers. DTT_{fin} was lower for the experienced drivers, and the stopping consistency was higher (lower SD DTT_{fin}), as could be expected. BrakeEvents was higher for experienced drivers, and Max. Brake was lower.

b) The adaptation analysis showed that during the driving session, drivers started to brake later with similar speed, which was shown by the following measures: smaller DTT_{ini}, lower TTC_{ini}, lower StopTime, and StopDistance. Drivers increased their stopping consistency (lower SD DTT_{fin}) and had fewer fluctuations in their deceleration profile, which was shown by: lower Dec. Inversions, higher R^2, and lower Max. Throttle.

c) The correlation matrix revealed that the braking maneuver is largely determined by the distance to the target at braking onset (DTT_{ini}) and the closely related TTC_{ini} (closely related because of the conditioned speed in this experiment). These correlate with many other measures which are related to the overall characteristics of the braking maneuver, such as Max. dec., Onset jerk, Max. Brake, StopTime, and StopDistance. The accuracy with which the vehicle is placed near the target can be expressed by two relatively uncorrelated accuracy measures: DTT_{fin} and SD DTT_{fin}.

d) The split-half reliability analysis shows that the scores on the measures Mean V, V_{ini}, DTT_{fin}, Max. dec., Onset jerk, TTC_{min}, SD dec., and Max. Brake have a high reliability (correlation coefficient between odd and even stops > .65). The lowest reliability was found for the stopping consistency measure (SD DTT_{fin}). The low reliability can be explained by the fact that this measure is based on a standard deviation amongst stops instead of a mean amongst stops, therefore requiring a larger sample size for high reliability.

1.4 Conclusions

The simulations provided insights into the characteristics of three complex braking measures. The R^2_{time} measure was suggested to replace $R^2_{distance}$, because the former distinguishes better between constant deceleration and multimodal braking. Furthermore, the simulations showed that the TTC_{min} measure is sensitive to distance offsets. The necessity to determine an (arbitrary) start and end-point for the mean TTC-dot calculations is a weak point of this measure, although it is successful in measuring an increase or decrease of the deceleration during braking.

The empirical study showed that the distance and time to the target at the onset of braking are the dominating variables determining the global characteristics of the braking maneuver.

A valid, reliable, and unique accuracy measure was provided by DTT_{fin}. This measure was successful in discriminating between experience levels. The stopping consistency (SD DTT_{fin}) performed likewise, but proved to be unreliable when too little stops are taken into account.

Table 4. Results of the experience comparison, adaptation analysis and split-half reliability.

	Experience comparison				Adaptation analysis				Split-half
	Inexp.	Exp.	*p*	*d*	First4	Last4	*p*	*d*	correlation
Nr. ValidStops	8.56	7.77	.13	0.39	3.23	3.35	.46	−0.11	.53
Mean V	10.16	10.12	.90	0.03	10.19	10.43	.27	−0.19	.70
V_{ini}	15.77	16.13	.54	−0.16	16.76	16.44	.41	0.13	.75
DTT_{ini}	47.84	53.38	.26	−0.30	59.66	50.45	.02	0.39	.48
DTT_{fin}	6.00	4.35	.04	0.54	5.61	5.07	.19	0.15	.69
SD DTT_{fin}	3.38	2.36	.05	0.55	3.07	2.09	.01	0.41	.09
Max. dec.	6.92	6.33	.17	0.36	6.77	6.66	.36	0.06	.85
Onset jerk	7.16	5.22	.08	0.48	6.00	6.27	.90	−0.05	.73
TTC_{ini}	2.94	3.10	.53	−0.17	3.44	2.91	.01	0.42	.55
TTC_{min}	1.55	1.54	.94	0.02	1.57	1.53	.62	0.07	.73
Max. dec. pos.	47.19	42.14	.14	0.41	48.11	45.13	.22	0.18	.41
Dec. inversions	2.92	3.32	.09	−0.45	3.45	3.01	.01	0.35	.29
SD dec.	2.09	1.85	.14	0.39	2.03	2.01	.65	0.02	.77
TransferTime	1.97	2.24	.29	−0.28	2.19	2.10	.86	0.07	.31
BrakeEvents	1.34	1.57	.04	−0.56	1.51	1.41	.22	0.18	.61
ThrottleEvents	0.18	0.19	.81	−0.06	0.25	0.16	.16	0.25	.33
Max. Brake	0.65	0.54	.03	0.58	0.61	0.60	.20	0.08	.84
Max. Throttle	0.22	0.19	.26	0.30	0.23	0.17	.01	0.49	.54
StopTime	5.07	6.14	.05	−0.52	6.28	5.42	.04	0.32	.63
StopDistance	41.84	49.03	.13	−0.41	54.05	45.38	.02	0.37	.43
R^2_{time}	0.93	0.93	.47	0.19	0.91	0.95	.02	−0.44	.40
mean TTC-dot	−0.47	−0.46	.35	−0.24	−0.47	−0.47	.97	0.01	.43

The empirical study used data from earlier experiments, in which speed was regulated by traffic signs. Participants were instructed to obey the traffic signs, and this might have been of influence on differences found, for example, on the speed at braking onset measure. With other experimental designs, other conclusions could have resulted concerning this and other measures. Furthermore, we only looked at adaptation and experience, not at, for example, the influence of systems feeding back accelerations. Literature shows that for such systems, the maximum deceleration and brake onset jerk often reveal significant differences. We did show that the maximum deceleration and onset jerk correlate with each other and with the overall brake maneuver characteristics determined by the speed and distance at brake onset. We advise that these measures should be included in an analysis with the note that they are intercorrelated. They provide information on *how* participants slowed down.

We suggest that the following measures should be included in braking maneuver analyses: speed at braking onset (V_{ini}), distance at braking onset (DTT_{ini}), the stopping position (DTT_{fin}) as accuracy measure, and the multi-modality of the braking (R^2_{time}). Recommended additional measures are: the mean speed over the complete braking maneuver (Mean V), stopping consistency (SD DTT_{fin}), maximum deceleration (Max. dec.), and Onset jerk.

Table 5. Correlation matrix.

	Nr. ValidStops	Mean V	V_{ini}	DTT_{ini}	DTT_{fin}	SD DTT_{fin}	Max. dec.	Onset jerk	TTC_{ini}	TTC_{min}	Max. dec. pos.	Dec. inversions	SD dec.	TransferTime	BrakeEvents	ThrottleEvents	Max. Brake	Max. Throttle	StopTime	StopDistance	R^2_{time}	mean TTC-dot
Nr. ValidStops																						
Mean V	.40																					
V_{ini}	.30	.76																				
DTT_{ini}	−.33	−.01	.36																			
DTT_{fin}	−.08	.02	.20	.26																		
SD DTT_{fin}	−.11	.11	.24	−.04	.33																	
Max. dec.	.06	.54	.34	−.06	−.01	.18																
Onset jerk	.07	.36	.18	−.58	.02	.28	.74															
TTC_{ini}	−.24	−.29	.04	.93	.24	−.08	−.70	−.67														
TTC_{min}	−.01	−.42	−.13	.53	.68	.10	−.60	−.39	.63													
Max. dec. pos.	−.12	−.05	−.11	.36	.01	−.14	−.23	−.55	.40	−.06												
Dec. inversions	−.12	.03	.05	.06	−.18	−.17	.16	−.12	.00	−.22	.16											
SD dec.	−.03	.67	.58	−.45	.12	.33	.92	.73	−.66	−.50	−.32	.03										
TransferTime	−.01	−.26	−.42	−.13	−.03	−.03	−.16	−.15	−.01	.10	.04	−.02	−.23									
BrakeEvents	−.17	−.27	.01	.55	.10	−.05	−.33	−.35	.59	.47	−.11	.20	−.36	.13								
ThrottleEvents	−.18	.08	.23	.45	−.14	−.09	04	−.08	.45	−.03	.06	.19	.03	.00	.41							
Max. Brake	.06	.49	.27	−.54	−.08	.17	.96	.72	−.66	−.64	−.13	.11	.87	−.08	−.35	.09						
Max. Throttle	−.18	.57	.59	.06	.20	.11	.50	.29	−.13	−.19	−.04	.04	.61	−.06	−.14	.24	.57					
StopTime	−.23	−.29	.06	.89	.05	−.17	−.72	−.63	.93	.09	.16	.01	−.66	.03	.71	.49	−.69	−.15				
StopDistance	−.33	−.01	.34	.99	.09	−.10	−.57	−.60	.92	.42	.37	.10	−.48	−.13	.55	.49	−.54	.02	.91			
R^2_{time}	−.10	.22	.14	−.10	.11	.14	−.17	−.05	−.20	−.04	.18	−.31	.00	−.09	−.51	−.61	−.19	−.13	−.28	−.12		
mean TTC-dot	−.01	−.24	−.04	.29	.34	−.19	−.50	−.09	.32	.63	−.29	−.13	−.36	.08	.21	−.08	−.55	−.20	.38	.24	.15	

Chapter 2. Non-vestibular motion cueing in a fixed-base driving simulator: effects on driver braking and cornering performance

Abstract

Motion platforms can be used to provide vestibular cues in a driving simulator, and have been shown to reduce driving speed and accelerations. However, motion platforms are expensive devices, and alternatives for providing motion cues need to be investigated. In independent experiments, the following eight low-cost non-vestibular motion cueing systems were tested by comparing driver performance to control groups driving without the cueing system: 1) tensioning seatbelt, 2) vibrating steering wheel, 3) motion seat, 4) screeching tire sound, 5) beeping sound, 6) road noise, 7) vibrating seat, and 8) pressure seat. Results showed that these systems are beneficial in reducing speed and accelerations and that they improve lane-keeping and/or stopping accuracy. Particularly the tensioning seatbelt system had a large influence on driver braking performance. This system reduced driving speed, increased stopping distance, reduced maximum deceleration, and increased stopping accuracy. It is concluded that low-cost non-vestibular motion cueing may be a welcome alternative for improving in-simulator performance such that it better matches real-world driving performance.

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

Parts of the results in this paper have been presented at earlier conferences (Boschloo, Wieringa, Kuipers, De Winter, & Mulder, 2005; De Winter, De Groot, Mulder, & Wieringa, 2007; De Winter, De Groot, Mulder, Wieringa, & Dankelman, 2008).

2.1 Introduction

Driving simulators are broadly used for research, training, and assessment. The effectiveness of a simulator depends to a large extent on its fidelity, or level of realism. A distinction can be made between two types of fidelity: objective (or: physical) fidelity and perceptual (or: psychological) fidelity (Advisory Group for Aerospace Research and Development, 1980; Bürki-Cohen, Soja, & Longridge, 1998). Objective fidelity is the extent to which the simulator replicates the physical characteristics of the simulated vehicle and environment, for example in terms of brightness and contrast of the visual display, or temporal synchronization of the physical motion. Perceptual fidelity—arguably a more valid criterion than objective fidelity—is defined as the degree to which the operator's performance and control strategies in the simulator and real vehicle correspond, as well as the degree to which the operator subjectively perceives the simulator to produce its real life counterpart. Researchers have been concerned with gaining an in-depth understanding of the relationship between user performance and perception in virtual environments in comparison to its real world counterpart (e.g., Bella, 2008; Kemeny & Panerai, 2003; Mania, Troscianko, Hawkes, & Chalmers, 2003; Nikooyan & Zadpoor, 2009; Shechtman, Classen, Awadzi, & Mann, 2009; Zaal, Pool, Mulder, Van Paassen, & Mulder, 2010). This study is also concerned with perceptual fidelity. More precisely, we will investigate means to make the simulator more realistic in terms of improving recorded driver performance and subjective driver experience.

A large share of the fidelity of a driving simulator is traditionally attributed to the implementation of a motion platform and the quality of its motion cues. Motion platforms feed back vehicle movements and accelerations, by tilting and translating the driver, and have been shown to be successful in reducing vehicle speeds and accelerations during many driving tasks. More specifically, motion platforms have been shown to result in lower onset jerk during braking, lower maximum decelerations during braking, lower cornering accelerations, better lateral vehicle control, more precise positioning of the vehicle to a stopping marker, and a higher subjective realism of the simulator (Brünger-Koch, Briest, & Vollrath, 2006; Colombet et al., 2008; Greenberg, Artz, Cathey, 2003; Pinto, Cavallo, Ohlmann, Espié, Rogé, 2004; Reymond, Kemeny, Droulez, & Berthoz, 2001; Siegler, Reymond, Kemeny, & Berthoz, 2001). With the advent of inexpensive high-end outside visual systems, motion platforms are becoming relatively more expensive. In many research applications as well as in commercial driver training, the implementation of a motion platform can seldom be justified, as a good motion platform often costs much more than a real car (Evans, 2004).

Fixed-base (i.e., without a motion platform) driving simulators avoid the cost issue. Literature shows that these stationary simulators provide driving performance measure scores which correlate well with those obtained in real-world driving, such as driving test performance and accident-involvement (e.g., Allen, Park, Cook, & Fiorentino, 2009; De Winter et al., 2009; Hoffman & McDowd, 2010; Lee, 2003). Although this *relative* validity is very encouraging, the *absolute* values of driving speeds, accelerations, and number of driving errors are generally considerably higher in fixed-base simulators than in reality (Boer, Girshik, Yamamura, & Kuge, 2000; Green, 2005; Hurwitz, Knodler, & Dulaski, 2005; Reed & Green, 1999). This impairs the validity of the research conducted with these devices and likely hampers a fixed base simulator's training effectiveness. Whereas normal driving on the road typically results in sustained accelerations of 4 m/s^2 during cornering or braking, decelerations of 6 or 7 m/s^2 are not uncommon in fixed-base simulators (cf. Siegler et al., 2001). Such deviant performance is often attributed to the lack of physical motion cueing, leaving the driver with only the visual system to perceive one's locomotion through the environment, which in turn results in inferior speed perception as compared to reality (Boer et al., 2000; Greenberg et al., 2003). In conclusion, low-cost alternatives for providing motion cues need to be investigated.

2.1.1 Non-vestibular motion cueing

In the context of aviation, Vaden and Hall (2005), as well as Bürki-Cohen, Sparko, and Go (2007) commented that a motion chair may offer much of the advantages of motion platforms without the disadvantage of high costs. With alternative motion cues, which are not (primarily) aimed at stimulating the vestibular organs, there is a potential to provide fully proportional and sustained acceleration feedback. For fighter aircraft, dynamic seats have been studied extensively and have been shown to yield a positive effect on pilot flying performance and the reported realism of the simulation (Ashworth, McKissick, & Parrish, 1984; Chung, Perry Jr., & Bengford, 2001; Flach, Riccio, McMillian, & Warren, 1986; Martin, 1986; Parrish & Steinmetz, 1983; Rutten, 1999; but see Showalter & Parris, 1980, for an example of no effect).

Riecke, Schulte-Pelkum, Caniard, and Bülthoff (2005) conducted experiments to achieve the illusion of self-motion in virtual reality without physically moving the observer. They investigated the effects of scene consistency, minor modifications of the projection screen, multi-sensory cue integration using seat vibrations, and auditory cues. Riecke et al. showed that the illusion of motion can be facilitated using these modifications, that is, without physical movement. Mollenhauer, Romano, and Brumm (2004) studied different types of motion that were presented by a motion seat in a driving simulator. They found that a motion seat without visual compensation of the seat's movements was preferred by the participants.

2.1.2 Research aim

This study aims to investigate whether low-cost non-vestibular motion cueing devices can be used to improve in-simulator driving performance. Our emphasis lies on reducing driving speed and accelerations in order to improve lane keeping and stopping precision in the simulator, and to obtain more realistic values that are comparable to those reported in typical real-world driving tasks. We describe eight independent experiments, in each of which the effects of a low-cost non-vestibular motion cueing systems was investigated using elementary braking and cornering tasks. In addition, a meta-analysis was conducted on the results of the eight experiments, in order to detect underlying regularities in the individual experiments. In the meta-analysis, the moderating role of driver experience was investigated as well.

2.2 Experiments

2.2.1 Apparatus

All experiments were conducted in a fixed-base driving simulator, called the Dutch Driving Simulator (Green Dino, 2008). The steering wheel, pedals, gear lever, ignition key, and seat resembled those of a real car, and the dashboard, interior, and mirrors were integrated in the projected outside world image as shown in Figure 1. Steering wheel force feedback was provided through an electrical motor according to the self-aligning torque of the front wheels. The simulator provided a horizontal visual field of view of 180 degrees through three projectors. The front view projection had a resolution of 1,024 x 786 pixels; the side-views featured resolutions of 800 x 600 pixels. The simulator provided realistic engine and wind sound from four speakers in the simulator cabin. The simulation ran at a frequency of 100 Hz, and the update rate of the visual projection was always larger than 25 Hz.

2.2.2 Experimental protocols

Eight experiments were performed independently during the course of four years, from 2005 to 2008. Each of these experiments evaluated a particular motion feedback device, and compared driving performance as obtained with the system enabled versus disabled (i.e., motion On vs. Off), using either a between-subjects or a within-subjects experimental design. The high similarity between the experimental protocols allows for a joint investigation of the *relative* effects of each motion system. Comparison between experiments of the *absolute* effects can only be done with care, however, because inevitably the experimental protocols differed from each other. Table 1 provides relevant details about the experimental protocols.

In the between-subjects experiments, the participants were allocated at random or alternately to either the motion On or Off group. The same method was applied for the within-subjects-experiments, with the additional procedure that participants who drove with motion On in a session drove with motion Off in the following session, and vice versa. In one of the between-subject experiments (motion seat) we deviated from the above procedure because of practical difficulties of changing the motion conditions. In this experiment, the participants drove with the same motion condition on a single day. However, we considered it unlikely that unidentified sequence effects contaminated the results, because the effects of the braking experiment were replicated in an additional within-subject experiment with the motion seat (De Winter, De Groot, Mulder, & Wieringa, 2007; N = 24; data not included in this study).

Figure 1. Driving simulator as used for the experiments.

Table 1. Experimental protocols.

	Number of participants	Number of women	Experienced/ Inexperienced*	Within / Between subject design	Number of sessions per participant	Number of maneuvers**	Task***	Automated controls****
1. Tensioning seatbelt	20	0	I	W	2	10 stops	A	S+G
2. Vibrating wheel	13	2	E	W	2	10 stops	A	-
3. Motion seat	60	0	E	B	1	10 stops	A	G
3. Motion seat					1	4 turns	C	G
4. Screeching tires	12	0	I	W	4	4 stops	A	S+G
4. Screeching tires					4	4 turns	B	G
5. Beeping sound	28	0	I	B	1	6 stops	A	S+G
5. Beeping sound					2	4 turns	B	G
5. Beeping sound					1	14 turns	C	G
6. Road noise	36	9	I	B	1	10 stops	A	S+G
6. Road noise					2	14 turns	C	G
7. Vibrating seat	15	3	E	W	4	8 turns	B	-
8. Pressure seat	31	5	I	B	1	14 turns	C	G

*Experienced is defined as being in the possession of a driver's license.
**Number of maneuvers indicates how many maneuvers of one driving session were taken into account in the statistical analyses.
***Task indicates the task that had to be performed. Three different tasks were used, as explained in the text.
****Indicates whether steering (S) and/or gear changing (G) were automated.

The experiments consisted of one or more of three different tasks marked as Task A, Task B, and Task C. All tasks were conducted without any other traffic present in the simulated world. During all experiments, participants had 4 to 5 minutes of driving time in a different virtual world, in order to get familiarized with the simulator before the actual experiment started. The speed limit was indicated by road signs during Task A and Task B, and during Task C pre-session instructions about the speed limit were given. Participants were given no further feedback about their driving speed.

During Task A, participants had to drive along a straight road containing many intersections, which were marked with a stop sign and a stop line. The speed limit varied from intersection to intersection between 30, 50, and 80 km/h, as indicated by road signs. Braking performance was analyzed during task A, and the stops at intersections 2 to 11 were included in the analysis per driving session. Task B was a cornering task during which participants had to turn either left or right at intersections. These intersections had no specific road markers, and without the presence of other vehicles, participants were free to choose their path and speed through the turn. The turn direction was kept constant during one driving session, thereby essentially driving "a square around the block". After the first session participants had to perform the same task with motion in the opposite direction in the second driving session. Turns 2 to 5 of every driving session were included in the analysis. The speed limit during task B was 50 km/h. Task C was a 7.5 km closed two-lane lap, consisting of 25 straight road segments, and 25 curves (21 of which were 90 degree curves). The lap contained no intersections, and a speed limit of 80 km/h was in effect. Participants were asked to drive around in this world and follow the road. The first fourteen 90 degree curves were used in the data analysis. The lane width was 5 m.

2.2.3. Participants and instructions to participants

All participants were recruited from the Delft University of Technology community, and were almost all undergraduate students at the Mechanical Engineering faculty. Participants were not informed about the precise purpose of the experiment, and instructions were kept to a minimum. Participants were just asked to perform a driving task or complete a driving lesson. For a number of experiments experienced participants were used, the other experiments featured inexperienced drivers. Experience was defined as “having a driver’s license.” When experienced drivers were used, they were asked to drive as they normally would while respecting the traffic laws. When inexperienced drivers were used, they were given the minimum information necessary to complete a (simplified) driving task. Most experiments were completed with an automatic gearbox; some braking experiments were completed with automated steering as well as an automatic gearbox. Table 1 summarizes this information for all experiments.

2.2.4 Dependent measures and statistical analyses

Tables 2 and 3 summarize the dependent measures that were used for the analyses. The measures were largely based on the work of Siegler et al. (2001), as they analyzed motion versus no motion during braking and cornering. For the stopping task (Task A) speed and distance at the onset of braking (V_{ini} and DTT_{ini}) are important, since they effectively determine how hard the driver should brake in order to come to a stop at the stop line. The final stopping position (DTT_{fin}) is a measure of precision, and the R^2_{time} measure describes how drivers slowed down, that is, in one movement or in various stages with multiple brake pedal modulations. These four measures have been shown to be valid and unique descriptors of braking performance (De Groot, De Winter, Wieringa, & Mulder, 2009). Additionally, the mean speed over the complete stopping maneuver, the stopping consistency (SD DTT_{fin}, measured as the standard deviation of all distances-to-stop-target in the session), the maximum deceleration (Max. dec.), and the initial rate of change of deceleration (Onset jerk) provide further information about braking performance. Note that the braking variables did not include the foot-off-gas-pedal time prior to pressing the brake; the onset of braking was defined as the first moment of pressing of the brake pedal when being less than 175 m before the intersection.

For the cornering tasks (Tasks B and C), the measures LCE45deg (Lane Center Error at the midpoint of the 90 degree turn) and V45deg (Velocity at the midpoint of the 90 degree turn) provide information on the path and speed halfway through the turns. Because only relatively small changes in vehicle paths can be expected, the speed effectively determines the lateral acceleration required to complete a turn. We therefore decided not to include the lateral acceleration in the measures, also because of the large variance of the lateral acceleration during cornering. Siegler et al. (2001) showed that the LCE45deg and V45deg measures were sensitive measures to describe cornering behavior, as opposed to the lateral acceleration at the midpoint. Finally, the SDLP (Standard Deviation of the Lateral Position) measure was calculated for the complete turn, which describes the variance of the vehicle path compared to the average lateral position. For corners without a clear center-line (i.e., during Task B), the lateral position was calculated as compared to a participant’s personal mean path (resulting in the SDLP measure), as opposed to the mean lateral position with respect to the road's center-line. For these measures a distinction was made between left (L) and right (R) corners, as left and right corners have different radii, lane markings, and lane boundaries, and large differences in behavior between left and right turns can be expected.

Besides the objective measures, a short questionnaire addressing the participants' subjective experience of the simulator's realism and/or the participant's own driving performance in the simulator was used in each experiment. If a questionnaire item contained the words realism or realistic (e.g., "The sound in the simulator is realistic"), then this item was included in the calculation of a subjective realism score. For some experiments only one question was used to calculate a participant's score, for other experiments the average of five questions was used. All questionnaires contained 4-point, 5-point, or 10-point Likert items, except the screeching tires experiment in which only open questions were asked, and the vibrating seat in which binary answers could be given. The Likert scale data was transformed to a percentage ranging from 0 to 100 for the 4 to 10 entry possibilities and treated as interval scale data during further analysis.

The dependent measure scores and questionnaire data were tested for the effects of the experimental conditions using either a paired or an independent Student's *t* test, depending on whether the experiment was of within-subject or be-

Table 2. Dependent measures for Task A.

Abbreviation	Unit	Description
V_{ini}	m/s	Speed at onset of braking
DTT_{ini}	m	Distance to the target line at onset of braking
DTT_{fin}	m	Distance to the target line of the stopping position
R^2_{time}	-	Squared correlation coefficient of the speed vs. time data from brake onset to stop
Mean V	m/s	Mean speed from 175 m before to 30 m after the stop
SD DTT_{fin}	m	Session standard deviation of distance to the target line at stop position
Max. dec.	m/s^2	Maximum deceleration during braking for speed greater than 5 km/h ($t = T$ s)
Onset jerk	m/s^3	Mean initial rate of deceleration (deceleration at t=T/2 divided by T/2)

All measures were calculated for each participant over the 2nd to the 11th stop of the session; *t* represents the elapsed time since the onset of braking.

Table 3. Dependent measures for Tasks B and C.

Abbreviation	Unit	Description
NrDepartures*	#	Number of road departures
LCE45deg (L)	m	Lane center error when halfway through left turns
LCE45deg (R)	m	Lane center error when halfway through right turns
V45deg (L)	m/s	Speed when halfway through left turns
V45deg (R)	m/s	Speed when halfway through right turns
SDLP (L)	m	Standard deviation of the lateral position through left turns (task C)
SDLP (R)	m	Standard deviation of the lateral position through right turns (task C)
SDLP (L)	m	Standard deviation of lateral deviation from own mean path through left turns (task B)
SDLP (R)	m	Standard deviation of lateral deviation from own mean path through right turns (task B)

* All other measures during Task B were calculated for each participant as the mean of the second to the fifth turn of the session. For Task C, the dependent measures were calculated as the means of all 90 degree turns with radii of 15–20m. Data from 10 seconds prior to 20 seconds after road departures were removed.

tween-subject design. Note that the *t* test is a powerful and robust alternative to nonparametric tests, also for Likert-type questionnaire data (De Winter & Dodou, 2010). In the meta-analysis we used a combined independent-paired *t* test (Looney & Jones, 2003). Two-tailed *p*-values were calculated, as well as the 95% confidence interval for the true difference of population means. Cohen's *d* (i.e., the mean standardized difference) was used as an effect size measure; this standardized effect size allows for comparisons across studies and across the dependent measures.

2.2.5. The eight systems under evaluation

Eight motion cueing systems, all aimed at reducing driving speed and accelerations, were developed and evaluated: 1) tensioning seatbelt, 2) vibrating steering wheel, 3) motion seat, 4) screeching tire sound, 5) beeping sound, 6) road noise, 7) vibrating seat, and 8) pressure seat. The first two systems provided feedback of the car's longitudinal accelerations; systems (3) to (5) fed back both longitudinal as well as lateral accelerations; system (6) provided feedback on the car's speed, and systems (7) and (8) provided feedback of the car's lateral acceleration. The characteristics of all systems and experiments will be briefly described below.

1. Tensioning seatbelt (longitudinal acceleration cueing)

The feedback cue was a tension force in the seatbelt, proportional to the deceleration of the vehicle. Seatbelt tensioning has previously been used in aircraft simulation for lateral and vertical acceleration cueing by tensioning as well as laterally scrubbing the seatbelt over the pilot's lap (Heintzman, 1996). In the present study, the tension force of a standard seatbelt over the driver's left shoulder increased linearly from 0 to 150 N for decelerations between 0 and 5 m/s^2. For decelerations larger than 5 m/s^2, the force remained at 150 N. These force settings were based on the results of a just-noticeable-difference experiment (*N* = 8, data not shown), which showed that participants could not notice a force increase of 25 N, while a reduction of 25 N was noticed clearly. Increases of 50 N and above were clearly noticed by all participants.
The maximum force of 150 N was determined by the experimenters as the maximal force which could still be considered comfortable. To provide some sense of force magnitude, 150 N is approximately the force required to support the upper torso (assumed mass of 30 kg) during 0.5 g decelerations. During crashes, forces exceed-

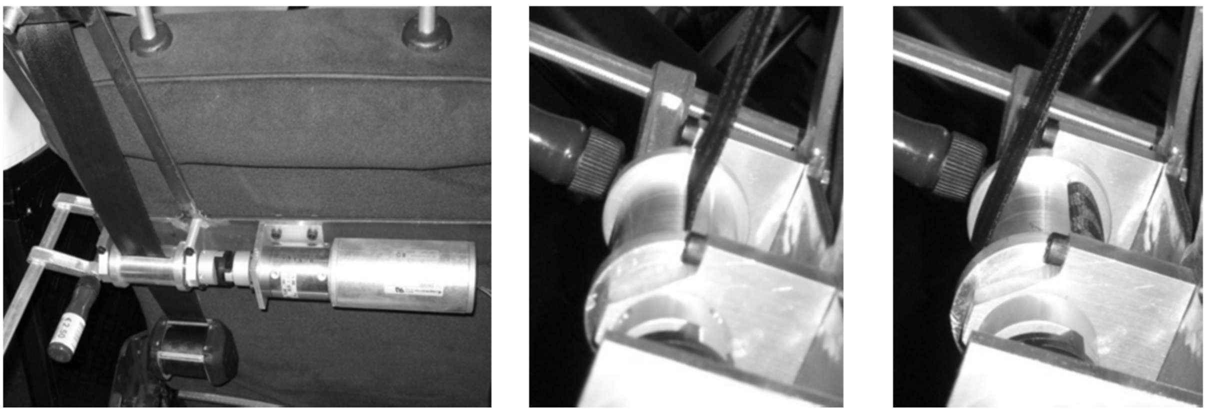

Figure 2. Tensioning seatbelt system on the back of the seat. The left picture shows the total assembly of motor, gearbox, coupling, and sleeved cylinder. The original belt tension device is also visible below the sleeved cylinder. The middle and right puctures show the sleeved cylinder in rest (middle),and with applied tensioning moment, with tensioned seatbelt (right).

ing 5,000 N can result (Gavelin, Lindquist, & Oldenburg, 2007). The seatbelt tension was generated by a moment-inducing motor with a reduction gearbox, as shown in Figure 2a. The seatbelt was fed through a sleeve in a cylinder on the back of the seat, and when a moment was generated by the motor, the belt was gripped by the cylinder, pulling the belt tightly over the shoulder and chest of the driver. This effect is illustrated by Figures 2b and 2c. When no moment was exerted, the seatbelt could be operated in regular fashion. The experiment was run as a within-subject design, with 20 participants.

2. Vibrating steering wheel (longitudinal acceleration cueing)
Vibrating elements are relatively inexpensive and easy to implement. Transferring information through vibrations has already been successfully applied during driving, for example for presenting spatial warning signals (Ho, Reed, & Spence, 2006), lane departure warnings (Suzuki & Jansson, 2003), and also for navigation (Van Erp, 2005). Drivers have been shown to respond quickly and intuitively to vibrations (Suzuki & Jansson, 2003). During flight simulation, vibrations are used to present cues like stall onset, high speed buffet, landing gear extension, engine vibrations, turbulence, and runway roughness (Heintzman, 1996). In the present study, vibrations were provided by a bass speaker attached to the steering wheel, as shown in Figure 3. A low-frequency (75 Hz) sample was played. The volume of the sample increased with increasing deceleration of the car, according to the graph shown in Figure 4. The experimenters set up the vibrations so that the vibrations were detectable from small deceleration onwards, and set the upper limit of the volume so that the speaker did not produce audible or resonating sounds. The experiment was run with 13 participants as a within-subject design.

3. Motion seat (longitudinal and lateral acceleration cueing)
Mollenhauer et al. (2004) showed that a motion seat had a positive effect on driving performance and the subjective experience of realism, irrespectively of the motion tuning parameters. Bürki-Cohen et al. (2007) showed that a simulator with a motion seat may be useful for training airline pilots. In fighter jet simulation, motion seats, such as the ALCOGS (Advanced Low Cost g-cueing System, Heintzman, 1996; see also Flach et al., 1986; Martin, 1986), have been used to feed back information about

Figure 3. Bass speaker as fitted on the steering wheel, displayed without steering wheel cover.

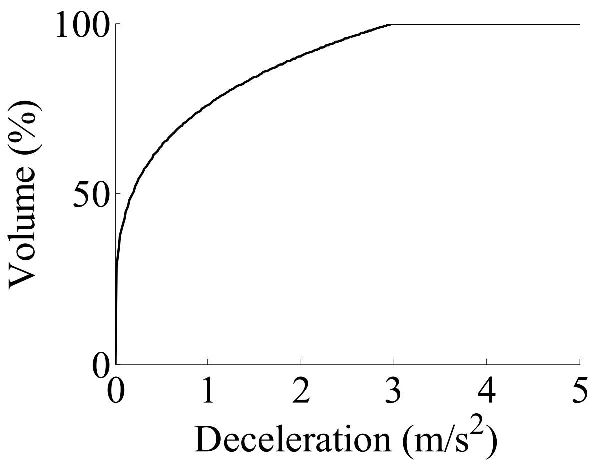

Figure 4. Volume regulation for increasing deceleration during the vibrating steering wheel experiment.

the state of the aircraft. In these two latter studies it was shown that a moving seat-pan can be used to provide vehicle roll or roll velocity information, which both improved performance up to the same level as of full body motion cueing. Other experiments showed that dynamic seats improve performance in compensatory and pursuit tracking in fighter aircraft simulators (Ashworth et al., 1984; Rutten, 1999). In our experiment, we used the motion seat shown by Figure 5 (Frex Japan Trading, 2008) during a braking and a cornering task. During braking, the seat tilted forwards proportionally with increasing deceleration until an angle of 4.7 degrees at 7.7 m/s^2. At higher decelerations the seat angle remained constant at 4.7 degrees. Two conditions were tested during the braking task: On and Off. During the cornering Task C, two motion cueing algorithms were tested versus the no motion condition (Off). During the normal seat movement, called the engineering approach (Eng), the driver's body is tilted outward in the turns, and a lateral acceleration of 8.3 m/s^2 corresponded to a seat inclination of 6.2 degrees. The second motion condition was exactly the opposite of Eng, and called the fun ride approach (Fun); the driver's body was tilted inwards in the turns. These motion conditions were inspired by a paper of Von der Heyde and Riecke (2001), in which they hypothesized that the fun ride approach would yield higher pleasure ratings and lower realism ratings than the engineering approach, and that no motion would be worse than motion on all measures.

The maximum rotation rate of the seat was 100 deg/s, and the visual scene was not compensated for the seat motions. This means that the seat moved with respect to the simulator cabin, and the visual projection as well as the pedals and steering wheel did not move at all. The experiment was a between-subject design using 60 participants. Two tasks had to be performed by the participants, firstly Task A, followed by Task C, always in this order.

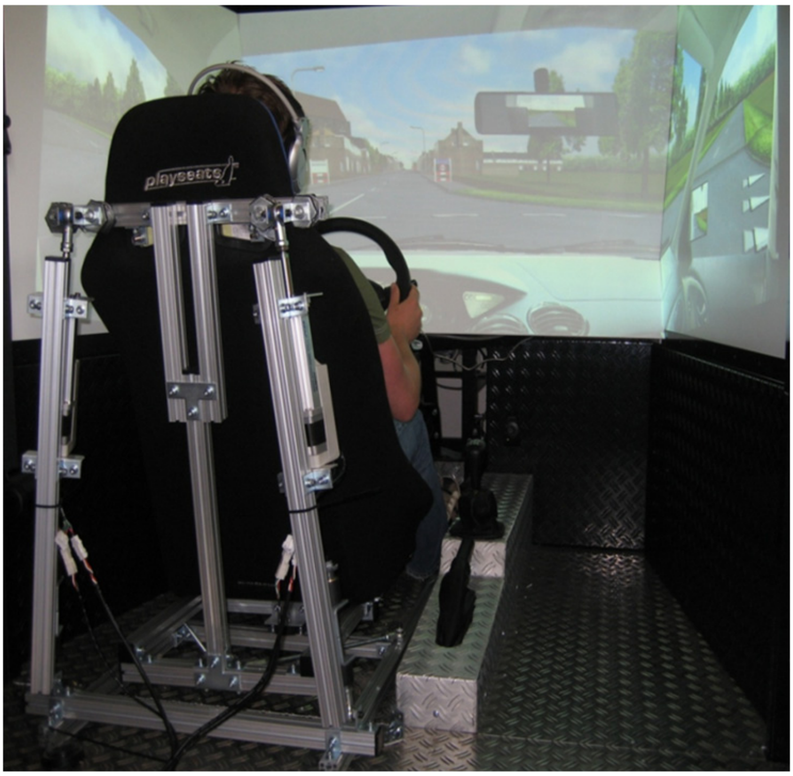

Figure 5. Motion seat positioned in the Dutch Driving Simulator. Note the two linear actuators on the back of the seat.

4. Screeching tire sound (longitudinal and lateral acceleration cueing)
Vehicle sound can increase the overall sensation of speed (Davis & Green, 1995). Not all simulators generate screeching tire sounds when driving near the performance limit of the tires. Considering the fact that people in simulators often drive unrealistically fast, providing screeching tire sounds seems a relevant cue to enhance driver awareness. The screeching sound was generated when the acceleration of the simulated car exceeded a friction ellipse (Milliken & Milliken, 1995) with a semi-major axis of 8.0 m/s^2 in longitudinal direction and a semi-minor axis of 7.2 m/s^2 in lateral direction. The sound volume of the screeching sound increased proportionally for increasing acceleration beyond the ellipse, with maximum volume at the absolute limit of the car (9.4 m/s^2 in longitudinal and 8.5 m/s^2 in lateral direction). The sound itself was a high pitch sound (frequency content between 1.0 and 1.8 kHz), based on a real world on-board sample of screeching tires. The sound volume level was about equal to the engine and wind noise of the simulator near the ellipse, and could increase to the dominant sound audible near the acceleration limit. The sound was presented non-directionally through four speakers in the four corners of the simulator cabin. This experiment was a within-subject design with 12 participants. Participants first had to complete 4 sessions of Task A, during which the system was alternately turned On and Off, and subsequently had to perform Task B, with a similar protocol.

5. Beeping sound (longitudinal and lateral acceleration cueing)
Similar to the screeching tire sound experiment described above, this experiment aimed at feeding back vehicle accelerations with sound. The supposed advantage of the less realistic beeping sound feedback with respect to the screeching tire sound is that it also feeds back accelerations during normal driving, instead of only near the acceleration limit. Here, a beeping sound was used with a tone of 1.0 kHz and beep duration of 0.2 seconds. The beeping frequency depended on the magnitude of the vehicle's total acceleration vector, calculated as the square root of the sum of the squared longitudinal and lateral accelerations. For accelerations from 1 m/s^2 up to 7 m/s^2, a beeping frequency starting at 0.5 Hz and increasing up to 7 Hz, according to the graph shown in Figure 6, was presented. A dead-zone between 0 and 1 m/s^2 was implemented to avoid beeping when the vehicle was standing still or driving with constant velocity. The function shown in Figure 6 was implemented to maximize beeping frequency differences between accelerations of 2 and 5 m/s^2 while having

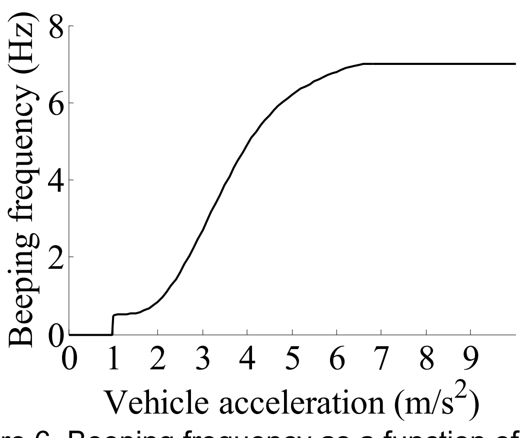

Figure 6. Beeping frequency as a function of total vehicle acceleration.

smooth transitions between frequencies. The sound volume level was designed so that it was clearly audible over the already present engine and wind noise. The sound was presented non-directionally through four speakers placed in the four corners of the simulator cabin. This experiment was a between subject design with 28 participants, and participants had to perform Tasks A, B and C consecutively. During Task A and Task B, the On group drove with the system enabled and the Off group with the system disabled, while the system was disabled for both groups during Task C, to test for a short term transfer of training effect.

6. Road noise (vehicle speed cueing)
Seat-base vibrations were used to simulate the effects of road noise. This system fed back vehicle speed, not accelerations like all other systems in this paper. The vibrations were provided by 26 small tactile display elements as found in mobile phones, which were equally distributed underneath the driver's legs (13 per side, as shown by Figures 7a and 7b). The tactile elements were controlled for speeds of 0 km/h up to 90 km/h. From 0 km/h, the excitation voltage of the display elements was increased as a quadratic function of speed up to the maximum excitation voltage at 90 km/h, from which the voltage remained constant with increasing speed. With increasing voltage, the rotational speed of the motor increases, and the eccentric mass causes an increase in vibration amplitude. Participants performed Task A, followed by two sessions of Task C. This experiment was a between-subject design with 36 participants. Two groups were made, one group drove all but the final session with road noise, and the other group always drove without road noise feedback. During the final session (Task C) all participants drove without feedback and the vibration group was tested for a short-term transfer of training effect.

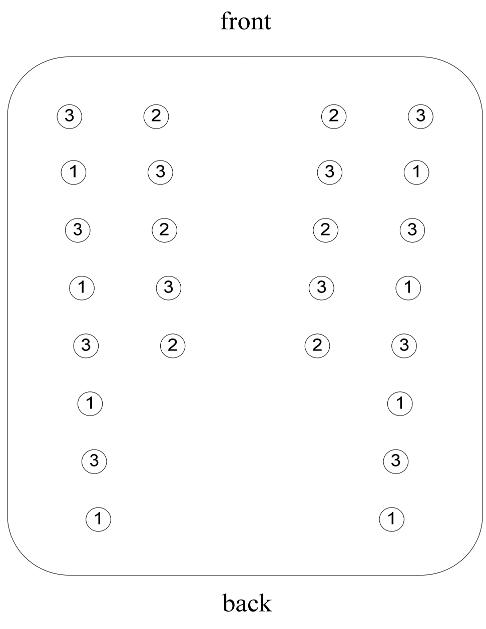

Figure 7a. Seat cover with 26 vibration elements. The elements are positioned under the driver's left and right legs.

Figure 7b. Seat cover as used in the simulator. The two most centrally placed actuators were not used.

7. Vibrating seat (lateral acceleration cueing)
To provide a sense of lateral acceleration to the driver, vibrations depending on vehicle lateral acceleration were provided by tactile display elements on the bottom of the seat, with the same hardware as was used in the speed-dependent road noise described above. When driving through right turns, elements in the left part of the seat vibrated and vice versa. For lateral accelerations above 1 m/s^2, four elements (numbered 1 in Figure 7a) vibrated on one side. Above 3 m/s^2, three additional elements (numbered 2 in Figure 7a) also started vibrating and above 5 m/s^2 all elements (numbers 1, 2, and 3 in Figure 7a) under the outside leg in the turn would vibrate. This provided participants with a stepwise proportional cueing. This experiment was a within-subject design with 15 participants. Participants had to complete 4 sessions of Task B, in which the system was enabled and disabled alternatively.

8. Pressure seat (lateral acceleration cueing)
When turning in a real car, the driver is pressed to the side of the seat, and as a result feels a reaction force on the side of his back. To simulate these reaction forces in a qualitatively correct fashion, the seat of the simulator was equipped with two pneumatic cylinders. For flying, similar pneumatic seats have been used, for example the NASA Langley Dynamic Seat (Heintzman, 1996). The general idea of these systems is that strong stimulation of selected parts (small surface, high pressure) of the body provides considerable feedback, while the reaction forces (which are equally large as the stimuli because the participant remains in the same position) are distributed over a larger part of the body (larger surface, lower pressure), providing less feedback. Here, pneumatic pressure was regulated proportionally to lateral accelerations. The actuators were fitted with metal plates, on the right and left side of the seat back as shown in Figures 8a and 8b. In left turns, the plate on the right side was pressurized and the plate on the left side remained unpressurized. In right turns, this situation was reversed. The maximal force of approximately 200 N was chosen by the experimenters as being the largest force which could still be considered comfortable while driving. This experiment was a between-subject design with 31 participants. Participants had to complete one driving session of Task C.

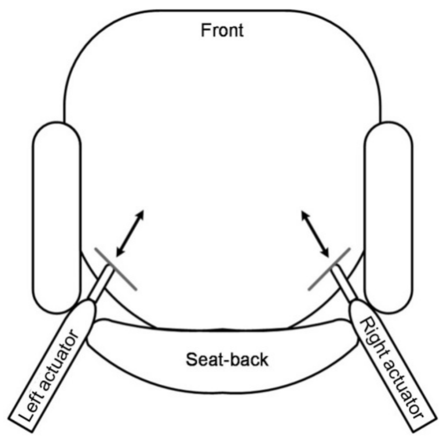

Figure 8a. Schematic picture of the pressure seat. Note the placing and orientation of the two pneumatic actuators. The right actuator was active in left turns and vice versa.

Figure 8b. Pressure system fitted in the back of the simulator's seat, with the two metal plates fitted to the pneumatic cylinders.

2.3 Results

The results of Task A (the braking task) are shown in Table 4; the results of Tasks B and C (cornering tasks) are shown in Table 5.

2.3.1. Tensioning seatbelt (longitudinal acceleration cueing)

The tensioning seatbelt system had a large effect on braking performance. As a result of the system, participants initiated braking at a larger distance from the intersection (p = .013, d = −0.69) and at lower speeds (p < .001, d = 0.55). They also stopped closer to the stop line (p = .016, d = 0.53) and slowed down with a smoother deceleration profile as measured by a higher R^2_{time} score (p = .026, d = −0.71). Additionally, the mean speed over the complete maneuver was lower (p = .005, d = 0.46), as were the maximum decelerations (p < .001, d = 0.87) and onset jerk (p < .001, d = 0.99). The stopping consistency was also better with the system enabled (p = .009, d = 0.52). Participants reported to have noticed the seatbelt very well (mean = 4.6 on a questionnaire item running from 1 (*poorly noticed*) to 5 (*strongly noticed*), 95% confidence interval 4.32–4.88).

2.3.2 Vibrating steering wheel (longitudinal acceleration cueing)

The vibrating steering wheel had a relatively small effect on braking performance. The only significant effect was that the vibrations resulted in a lower brake onset jerk, which is the derivative of acceleration at the start of the braking maneuver (p = .013, d = 0.54). It is interesting to note that other performance measures changed in the expected direction, but not significantly. For example, the speed at brake initiation was reduced (p = .267, d = 0.24), the smoothness of the deceleration profile was improved (p = .178, d = −0.45), and the maximum deceleration was lower (p = .253, d = 0.25) as a result of the vibration system.

2.3.3 Motion seat (longitudinal and lateral acceleration cueing)

The motion seat caused a reduction of onset jerk from 5.4 to 3.0 m/s^3 during Task A (p = .001, d = 0.94). It also improved other brake performance measures, although not statistically significant. The largest effect sizes next to the onset jerk were found for the stopping position (p = .086, d = 0.48) and the stopping consistency (p = .186, d = 0.37).

During the cornering Task C, the motion seat did not result in any significant differences between performance measures. However, some insignificant trends were observed. The motion seat caused a reduction of the speed through left turns, for both the fun (p = .188, d = 0.42) and engineering approach (p = .139, d = 0.48). Through right turns this effect was not observed. The driving precision was increased for the engineering approach through left turns, as both the lane center error (p = .515, d = 0.21) and standard deviation of the lateral position (p = .189, d = 0.42) were reduced. The fun ride approach did not show this trend.

2.3.4 Screeching tire sound (longitudinal and lateral acceleration cueing)

The screeching tire sound did not result in significant effects of the dependent measures for the braking maneuver. The largest effects were found for the stopping consistency (p = .201, d = 0.50), onset jerk (p = .277, d = 0.46), and the deceleration smoothness (p = .160, d = −0.45), all in favor of the On condition. A remarkable but insignificant finding is that because of the screeching tire sound, participants started braking from higher speeds (p = .524, d = −0.22) and at shorter distances to the inter-

Table 4. Means of the dependent measures for the braking experiments.

1. (A) Tensioning seatbelt within (n_{pairs}= 20)						
	Off	On	*p*	*d*	95% low	95% high
V_{ini} (m/s)	17.02	15.84	<.001	0.55	0.64	1.72
DTT_{ini} (m)	61.29	74.00	.013	−0.69	−22.39	−3.02
DTT_{fin} (m)	6.18	4.63	.016	0.53	0.32	2.79
R^2_{time} (-)	0.92	0.96	.026	−0.71	−0.06	0.00
Mean V (m/s)	0.99	9.37	.005	0.46	0.21	1.02
SD DTT_{fin} (m)	3.50	2.66	.009	0.52	0.23	1.45
Max. dec. (m/s^2)	6.25	4.51	<.001	0.87	1.03	2.45
Onset jerk (m/s^3)	5.76	2.32	<.001	0.99	2.04	4.83

2. (A) Vibrating wheel within (n_{pairs}= 13)						
	Off	On	*p*	*d*	95% low	95% high
V_{ini} (m/s)	15.38	14.98	.267	0.24	−0.35	1.15
DTT_{ini} (m)	38.91	40.08	.632	−0.14	−6.35	4.01
DTT_{fin} (m)	3.21	2.98	.570	0.15	−0.61	1.06
R^2_{time} (-)	0.91	0.92	.178	−0.45	−0.05	0.01
Mean V (m/s)	8.57	8.45	.169	0.18	−0.06	0.32
SD DTT_{fin} (m)	1.99	1.94	.859	0.05	−0.65	0.77
Max. dec. (m/s^2)	6.87	6.57	.253	0.25	−0.24	0.82
Onset jerk (m/s^3)	7.62	5.48	.013	0.54	0.53	3.76

3. (A) Motion seat between (n_{off} = 20, n_{on} = 40)						
	Off	On	*p*	*d*	95% low	95% high
V_{ini} (m/s)	16.35	15.87	.511	0.18	−0.98	1.94
DTT_{ini} (m)	51.46	53.20	.734	−0.09	−11.93	8.45
DTT_{fin} (m)	4.16	2.72	.086	0.48	−0.21	3.09
R^2_{time} (-)	0.93	0.94	.401	−0.23	−0.03	0.01
Mean V (m/s)	10.43	10.21	.492	0.19	−0.42	0.87
SD DTT_{fin}(m)	2.55	2.03	.186	0.37	−0.25	1.28
Max. dec. (m/s^2)	6.46	6.03	.317	0.28	−0.42	1.26
Onset jerk (m/s^3)	5.40	3.04	.001	0.94	0.98	3.75

4. (A) Screeching tires within (n_{pairs}= 12)						
	Off	On	*p*	*d*	95% low	95% high
V_{ini} (m/s)	13.86	14.10	.524	−0.22	−1.04	0.56
DTT_{ini} (m)	39.47	36.76	.298	0.25	−2.74	8.14
DTT_{fin} (m)	5.72	5.41	.703	0.14	−1.42	2.03
R^2_{time} (-)	0.93	0.95	.160	−0.45	−0.05	0.01
Mean V (m/s)	8.96	9.01	.667	−0.11	−0.28	0.18
SD DTT_{fin} (m)	2.47	1.77	.201	0.50	−0.43	1.81
Max. dec. (m/s^2)	6.67	6.44	.595	0.19	−0.69	1.14
Onset jerk (m/s^3)	8.44	5.57	.277	0.46	−2.65	8.39

5. (A) Beeping sound between (n_{off}= 13, n_{on}= 15)						
	Off	On	*p*	*d*	95% low	95% high
V_{ini} (m/s)	16.93	15.93	.255	0.44	−0.77	2.77
DTT_{ini} (m)	60.11	56.72	.572	0.22	−8.79	15.57
DTT_{fin} (m)	5.16	5.31	.845	−0.08	−1.80	1.48
R^2_{time} (-)	0.91	0.88	.148	0.56	−0.01	0.07
Mean V (m/s)	9.62	9.19	.181	0.52	−0.21	1.08
SD DTT_{fin} (m)	2.58	2.14	.359	0.35	−0.53	1.40
Max. dec. (m/s^2)	6.54	6.07	.327	0.38	−0.49	1.41
Onset jerk (m/s^3)	4.97	3.66	.212	0.48	−0.80	3.43

Table 4. (continued)

	6. (A) Road noise between (n_{off}= 18, n_{on} = 18)					
	Off	On	*p*	*d*	95% low	95% high
V_{ini} (m/s)	15.20	14.43	.181	0.46	−0.37	1.91
DTT_{ini} (m)	48.61	52.43	.590	−0.18	−18.11	10.46
DTT_{fin} (m)	5.84	5.23	.519	0.22	−1.30	2.53
R^2_{time} (-)	0.93	0.93	.934	0.03	−0.03	0.04
Mean V (m/s)	10.08	9.43	.023	0.79	0.10	1.20
SD DTT_{fin} (m)	3.12	2.40	.350	0.32	−0.82	2.25
Max. dec. (m/s^2)	6.88	6.37	.367	0.30	−0.63	1.67
Onset jerk (m/s^3)	6.87	5.66	.495	0.23	−2.34	4.77

Table 5. Means of the dependent measures for the cornering experiments.

	3. (C) Motion seat between (n_{off}= 20, n_{on} = 20)					
	Off	Eng	*p*	*d*	95% low	95% high
NrDepartures (#)	0.70	0.50	.503	0.21	−0.40	0.80
LCE45deg (L) (m)	1.72	1.55	.515	0.21	−0.33	0.66
LCE45deg (R) (m)	−0.76	−0.77	.957	0.02	−0.26	0.27
V45deg (L) (m/s)	12.10	11.51	.139	0.48	−0.20	1.39
V45deg (R) (m/s)	9.94	9.72	.610	0.16	−0.65	1.10
SDLP (L) (m)	0.43	0.38	.189	0.42	−0.02	0.12
SDLP (R) (m)	0.54	0.52	.640	0.15	−0.09	0.14

	3. (C) Motion seat between (n_{off} = 20, n_{on}= 20)					
	Off	Fun	*p*	*d*	95% low	95% high
NrDepartures (#)	0.70	0.35	.195	0.42	−0.19	0.89
LCE45deg (L) (m)	1.72	1.72	.995	0.00	−0.59	0.59
LCE45deg (R) (m)	−0.76	−0.74	.882	−0.05	−0.26	0.22
V45deg (L) (m/s)	12.10	11.61	.188	0.42	−0.25	1.24
V45deg (R) (m/s)	9.94	10.01	.872	−0.05	−0.96	0.81
SDLP (L) (m)	0.43	0.42	.863	0.05	−0.08	0.09
SDLP (R) (m)	0.54	0.57	.731	−0.11	−0.16	0.11

	4. (B) Screeching tires within (n_{pairs} = 12)					
	Off	On	*p*	*d*	95% low	95% high
NrDepartures (#)	0.00	0.00	-	-	0.00	0.00
LCE45deg (L) (m)	1.36	0.92	.143	0.32	−0.17	1.05
LCE45deg (R) (m)	−0.89	−0.83	.388	−0.12	−0.21	0.09
V45deg (L) (m/s)	9.33	8.75	.075	0.46	−0.07	1.23
V45deg (R) (m/s)	8.32	8.51	.672	−0.16	−1.15	0.77
SDLP (L) (m)	0.41	0.31	.114	0.72	−0.03	0.24
SDLP (R) (m)	0.26	0.35	.184	−0.53	−0.25	0.05

	5. (B) Beeping sound between (n_{off} = 13, n_{on} = 15)					
	Off	On	*p*	*d*	95% low	95% high
NrDepartures (#)	0.23	0.07	.231	0.46	−0.11	0.44
LCE45deg (L) (m)	2.58	2.21	.585	0.21	−1.02	1.77
LCE45deg (R) (m)	−0.78	−0.36	.299	−0.40	−1.24	0.40
V45deg (L) (m/s)	10.83	9.82	.188	0.51	−0.53	2.55
V45deg (R) (m/s)	8.97	8.87	.878	0.06	−1.21	1.41
SDLP (L) (m)	0.61	0.55	.675	0.16	−0.22	0.33
SDLP (R) (m)	0.51	0.46	.743	0.13	−0.25	0.35

Table 5. (continued)

		5. (C) Beeping sound transfer (n_{off} = 13, n_{on} = 15)				
	Off	On → Off	*p*	*d*	95% low	95% high
NrDepartures (#)	0.23	0.00	.161	0.56	−0.10	0.56
LCE45deg (L) (m)	1.00	1.23	.453	−0.29	−0.84	0.39
LCE45deg (R) (m)	−0.70	−0.67	.830	−0.08	−0.33	0.27
V45deg (L) (m/s)	9.09	9.88	.268	−0.44	−2.21	0.64
V45deg (R) (m/s)	8.74	8.87	.806	−0.10	−1.29	1.01
SDLP (L) (m)	0.52	0.45	.214	0.49	−0.04	0.18
SDLP (R) (m)	0.56	0.49	.260	0.44	−0.06	0.21

		6. (C) Road noise between (n_{off} = 18, n_{on}= 18)				
	Off	On	*p*	*d*	95% low	95% high
NrDepartures (#)	1.22	1.11	.798	0.09	−0.77	0.99
LCE45deg (L) (m)	1.91	1.86	.909	0.04	−0.85	0.96
LCE45deg (R) (m)	−1.08	−0.78	.044	−0.70	−0.60	−0.01
V45deg (L) (m/s)	11.53	10.73	.165	0.47	−0.35	1.95
V45deg (R) (m/s)	10.67	10.13	.342	0.32	−0.60	1.68
SDLP (L) (m)	0.54	0.51	.568	0.19	−0.09	0.17
SDLP (R) (m)	0.91	0.73	.058	0.65	−0.01	0.36

		6. (C) Road noise transfer (n_{off}= 18, n_{on} = 18)				
	Off	On → Off	*p*	*d*	95% low	95% high
NrDepartures (#)	0.94	0.28	.042	0.70	0.02	1.31
LCE45deg (L) (m)	1.84	1.77	.841	0.07	−0.62	0.76
LCE45deg (R) (m)	−1.15	−1.05	.537	−0.21	−0.42	0.23
V45deg (L) (m/s)	11.98	11.90	.845	0.07	−0.78	0.94
V45deg (R) (m/s)	11.30	10.72	.279	0.37	−0.49	1.65
SDLP (L) (m)	0.58	0.52	.379	0.30	−0.07	0.18
SDLP (R) (m)	0.83	0.67	.131	0.52	−0.05	0.36

		7. (B) Vibrating seat within (n_{pairs} = 15)				
	Off	On	*p*	*d*	95% low	95% high
NrDepartures (#)	0.13	0.07	.334	0.22	−0.08	0.21
LCE45deg (L) (m)	1.50	1.53	.880	−0.02	−0.54	0.47
LCE45deg (R) (m)	−0.86	−0.75	.417	−0.18	−0.39	0.17
V45deg (L) (m/s)	9.10	9.35	.367	−0.16	−0.82	0.32
V45deg (R) (m/s)	7.98	7.33	.024	0.46	0.10	1.22
SDLP (L) (m)	0.38	0.39	.819	−0.07	−0.10	0.08
SDLP (R) (m)	0.30	0.23	.145	0.50	−0.03	0.17

		8. (C) Pressure seat between (n_{off} = 13, n_{on} = 18)				
	Off	On	*p*	*d*	95% low	95% high
NrDepartures (#)	1.62	1.22	.493	0.25	−0.76	1.55
LCE45deg (L) (m)	1.55	1.25	.333	0.36	−0.33	0.93
LCE45deg (R) (m)	−1.14	−0.78	.073	−0.68	−0.74	0.04
V45deg (L) (m/s)	10.79	10.46	.607	0.19	−0.94	1.58
V45deg (R) (m/s)	10.62	9.93	.328	0.36	−0.73	2.12
SDLP (L) (m)	0.58	0.52	.230	0.45	−0.04	0.17
SDLP (R) (m)	1.03	0.73	.003	1.16	0.11	0.50

p represents the probability of observing the given means by chance if the means would actually be equal; *d* represents Cohen's *d*, a measure of the size of the difference between the means in relation to the pooled standard deviation; 95% low and 95% high represent the 95% confidence interval for the true difference of population means.

section (p = .298, d = 0.25), but with slightly lower maximum decelerations (p = .595, d = 0.19). In this respect, it must be noted that the feedback was only presented when the longitudinal acceleration was larger than 8 m/s^2 in longitudinal and 7.2 m/s^2 in lateral direction. This feedback only appeared at the highest decelerations, and might cause participants to search for the onset limit of the sound, thereby increasing their average deceleration values.

An important aspect of the screeching tire sound is that it can act like a binary warning signal, leading to behavioral adaptation. The mean maximum deceleration for the combined On and Off groups was 7.1 m/s^2 in sessions 1 and 2 and 5.9 m/s^2 in sessions 3 and 4, which is a significant difference (p = .002, d = 0.60). An additional control group of five participants (data not shown) who drove without any screeching tire sounds showed constant maximum deceleration over time: 6.3 m/s^2 in sessions 1 and 2 and 6.2 m/s^2 in sessions 3 and 4 (p = .90, d = 0.03). So, whereas participants driving with screeching tire sound indeed decreased their maximum decelerations during later sessions, this effect was not found for participants driving without screeching tire sounds.

For the cornering Task B, no significant effects of the dependent measures were found. As with Task A, however, there were indications that participants adapted their performance due to the screeching tire sounds. The largest effects found were: a reduction of speed through left turns (p = .075, d = 0.46), an increase of driving precision through left turns (p = .14, d = 0.32), a reduction of the standard deviation of the lateral position through left turns (p = .114, d = 0.72) and an increase through right turns (p = .184, d = −0.53). The questionnaire revealed that the majority of participants who heard the screeching tires indicated that it positively influenced their driving performance, primarily as it led them to adopt lower cornering speeds, which is a desired behavioral adaptation.

2.3.5 Beeping sound (longitudinal and lateral acceleration cueing)

The beeping sound had no significant effects on braking performance. Some measures showed notable effect sizes: participants initiated braking with lower speeds (p = .255, d = 0.44), decelerated with a less constant deceleration (p = .148, d = 0.56), had a lower average speed over the braking maneuver (p = .181, d = 0.52), and a lower onset jerk (p = .212, d = 0.48).

During Task B there were no significant differences either, although there were some interesting effects. The speed through left turns was reduced from 10.8 to 9.8 m/s (p = .188, d = 0.51), while it remained equal on right turns (p = .878, d = 0.06). The lane center error was reduced particularly through right turns (p = .299, d = −0.40). Additionally, the beeping sound group had less road departures than the control group (p = .231, d = 0.46).

During the transfer session, consisting of Task C, in which both groups drove without the beeping system, the group formerly driving with beeping sound had less road departures (p = .161, d = 0.56) and smaller standard deviations of the lateral position through both left (p = .214, d = 0.49) and right (p = .260, d = 0.44) turns, which suggests that the Beeping sound group drove more precisely than the control group albeit statistically insignificant.

2.3.6 Road noise (vehicle speed cueing)

The road noise resulted in a lower mean speed during Task A (p = .023, d = 0.79), without significant or substantial effects on maximum deceleration or onset jerk. There was an insignificant reduction of speed at the onset of braking from 15.2 to

14.4 m/s (p = .181, d = 0.46). The effects on speed alone were expected as the feedback depended on vehicle speed and not acceleration.

For the first of the two Task C sessions, participants receiving speed feedback had a smaller lane center error through right turns (p = .044, d = −0.70). They also drove more slowly through left (p = .165, d = 0.47) as well as right (p = .342, d = 0.32) turns, although not significantly. The last notable effect was that the SDLP through right turns was smaller for the feedback group (p = .058, d = 0.65).

During the second cornering session (Task C with road noise feedback turned off for both groups) the group previously driving with road noise had less road departures (p = .042, d = 0.70). Interestingly, this group increased their speed compared to the session before, when they drove with feedback, but still had lower speeds and larger precision as compared to the group always driving without the vibration feedback.

2.3.7 Vibrating seat (lateral acceleration cueing)

The vibrating seat significantly reduced the speed halfway through right turns (p = .024, d = 0.46). The standard deviation of the lateral deviation from a person's own mean path through right turns was reduced from 0.30 to 0.23 m (p = .145, d = 0.50), this was not significant though. Another insignificant effect was a reduced number of road departures (p = .334, d = 0.22). The speed through left turns, however, as well as the distance to the road center was not different.

2.3.8 Pressure seat (lateral acceleration cueing)

Only the standard deviation of the lateral position through right turns was significantly smaller as a result of the pressure seat system (p = .003, d = 1.16). Through left turns the standard deviation was also smaller, but to a lesser extent (p = .230, d = 0.45). Lane center errors through left (p = .333, d = 0.36) and right (p = .073, d = −0.68) turns were reduced considerably and the speed through right turns was reduced as well (p = .328, d = 0.36), although none of these effects were significant at the 5% level. The questionnaire revealed a trend (not significant) that participants driving with the pressure seat felt *less* like they were driving in a real car, as compared to the participants driving without the system. It must be noted, however, that for this experiment inexperienced participants were used and no explanation regarding what is normally felt, or about the principle of the pressure seat, was given before the experiment. Some participants indicated afterwards that they had not understood that the pressure on their back corresponded to the lateral acceleration of the car.

2.3.9 Subjective realism score

After all experiments a questionnaire was employed to determine a realism score. The screeching tires experiment used only open questions, and the vibrating seat experiment used questions with only two response options, and therefore these experiments were excluded from the realism score calculation. The vibrating steering wheel experiment used a 4-point Likert scale, the tensioning seatbelt and pressure seat experiments used 5-point Likert scales, and the motion seat, beeping sound and road noise experiments used 10-point Likert scales. The Likert scale scores were transformed to a percentage score, ranging from 0 to 100% for the first to the last Likert scale item. For the six experiments with Likert scale questionnaires, a realism score per participant was obtained. In the within-subject experiments a single questionnaire was employed after the complete experiment, and so the questionnaire could not be used to identify the influence of individual systems on subjective realism.

The results are shown in Table 6 and Figure 9. It must be stressed that the scores were not based on the same questions for all experiments and therefore only the relative values and not the absolute values can be compared. The effect sizes per experiment are of particular interest. All systems, except for the pressure seat, induced an increase of the realism score. For the motion seat, both the engineering and fun ride approaches of motion cueing resulted in a significant increase of the realism score compared to no-motion feedback.

2.3.10 Meta-analysis of all braking and cornering experiments

The combined effect size of all measures for Tasks A, B, and C was determined. The results are presented in Table 7. The overall *d*-statistics were calculated by combining the data for each experiment while adjusting for the mean values. The *p*-values and 95% confidence intervals were calculated with an independent *t* tests. The results show that the overall largest effect of motion feedback on braking performance is a reduction of brake onset jerk ($p < .001$, $d = 0.57$). A significant effect was also found for the average speed across a stopping maneuver ($p = .002$, $d = 0.35$). Similar effect sizes were found for an improvement of stopping consistency

Table 6. Subjective realism scores for all experiments

	Off	On	*p*	*d*
1. Tensioning seatbelt	36.3		-	-
2. Vibrating steering wheel	60.3		-	-
3. Motion seat Engineering	36.1*	47.2	.027	−0.73
3. Motion seat Fun	36.1*	48.2	.026	−0.73
4. Screeching tires	-		-	-
5. Beeping sound	50.4	56.5	.308	−0.39
6. Road noise	48.1	52.5	.504	−0.22
7. Vibrating seat	-		-	-
8. Pressure seat	50.0	41.2	.241	0.44

*Two experiment results were tested against the same baseline score.
p represents the probability of observing the given means by chance if the means would actually be equal; *d* represents Cohen's *d*, a measure of the size of the difference between the means in relation to the pooled standard deviation.

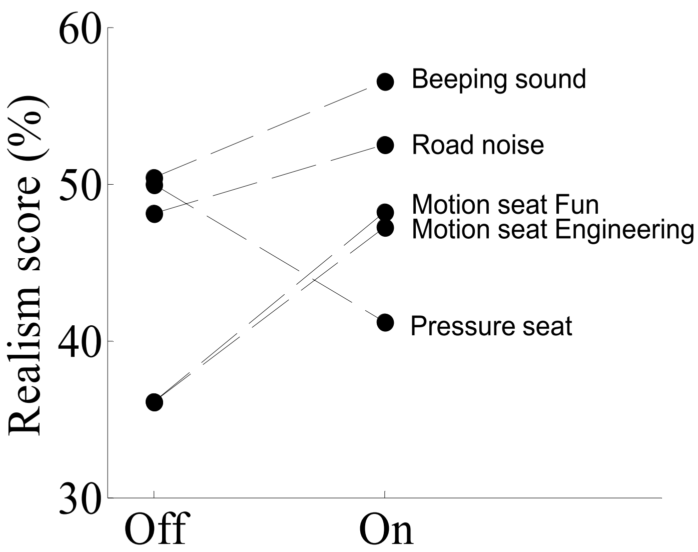

Figure 9. Realism scores for experiments with available questionnaire data for the two experimental conditions.

and reduction of maximum decelerations (d = 0.36 and 0.41, respectively). For the maximum decelerations during Task A, an illustration is presented by Figure 10, showing that all systems resulted in lower maximum decelerations. The feedback systems also caused drivers to stop closer to the stopping target (p = .015, d = 0.31) and initiate braking at a lower speed (p = .014, d = 0.30).

For cornering (Tasks B & C), the strongest effect was a reduction of speed through left turns (p = .011, d = 0.33). It is interesting to observe that this effect was not as strong for the speed through right turns. During right turns, the feedback systems caused a reduction in lane center error (p = .009, d = −0.32). Additionally, the number of road departures was reduced when the feedback systems were active (p = .161, d = 0.20), although this effect was not statistically significant. Another significant effect was a reduction of standard deviation of the lateral position during right turns (p = .022, d = 0.33) and an insignificant reduction during left turns (p = .114, d = 0.23).

2.3.11 Meta-analysis of the influence of participant experience

In order to explain a part of the differences in measure scores found between experiments, it is useful to investigate the effect of participant experience during Task A. During the seatbelt tensioning, motion seat, and road noise experiments, similar conditions were present for both experienced and inexperienced drivers. Only the data of a participant's first driving session was included in the analysis, and only if in that session any feedback system was disabled (baseline condition). In our research laboratory we conducted experiments which were similar in experimental protocol to the motion feedback experiments, but with control loadings, and because only the

Table 7. Meta-analysis results of the eight systems.
Task A (6 experiments, n_{off} = 51, n_{on} = 72, n_{pairs} = 45).

	p	d	95% low	95% high
V_{ini} (m/s)	.014	0.30	0.12	1.11
DTT_{ini} (m)	.178	−0.18	−7.31	1.35
DTT_{fin} (m)	.015	0.31	0.15	1.44
R^2_{time} (-)	.139	−0.20	−0.02	0.00
Mean V (m/s)	.002	0.35	0.12	0.56
SD DTT_{fin} (m)	.005	0.36	0.17	0.95
Max. dec. (m/s^2)	.001	0.41	0.27	0.99
Onset jerk (m/s^3)	<.001	0.57	1.21	3.21

Tasks B and C (6 experiments, n_{off} = 64, n_{on} = 91, n_{pairs} = 27).

	p	d	95% low	95% high
NrDepartures (#)	.161	0.20	−0.07	0.43
LCE45deg (L) (m)	.252	0.14	−0.12	0.46
LCE45deg (R) (m)	.009	−0.32	−0.33	−0.05
V45deg (L) (m/s)	.011	0.33	0.12	0.88
V45deg (R) (m/s)	.137	0.20	−0.10	0.70
SDLP (L) (m)	.114	0.23	−0.01	0.09
SDLP (R) (m)	.022	0.33	0.01	0.15

p represents the probability of observing the given means by chance if the means would actually be equal; d represents Cohen's d, a measure of the size of the difference between the means in relation to the pooled standard deviation; 95% low and 95% high represent the 95% confidence interval for the true difference of population means.

data of a participant's first session without a motion system is used, the data of a control loading experiment was included in this analysis as well. All sessions consisted of 10 stopping maneuvers in the same environment with an automatic gearbox and the same brake pedal configuration.

The above mentioned conditions were applicable to a total of 34 inexperienced and 26 experienced drivers. The results of the inexperienced and experienced participants are compared with each other in Table 8. From this table, one can conclude that experienced drivers drive more precise, because they stopped closer to the stop target position (p = .046, d = 0.53) and also had a higher stopping consistency (p = .045, d = 0.53). Although not significantly, the experienced drivers initiated braking at larger distances from the stopping target (p = .348, d = −0.25), and also braked with a lower onset jerk (p = .081, d = 0.46) and lower maximum deceleration (p = .174, d = 0.36).

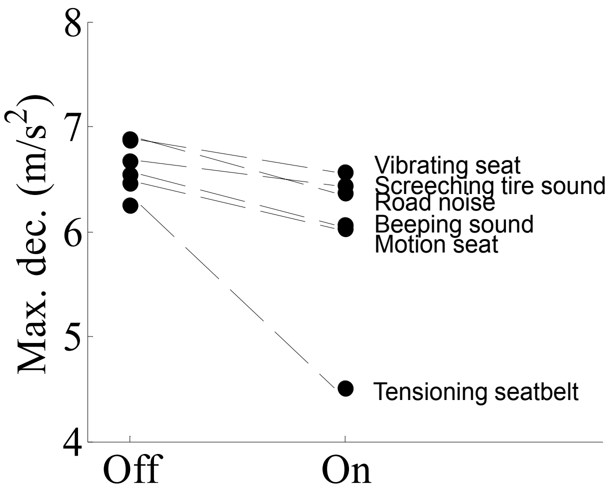

Figure 10. Mean maximum deceleration of brake experiments. Note the large effect of the tensioning seatbelt system as compared to the other systems.

Table 8. Experience comparison for Task A.

	Inexperienced (n = 34)	Experienced (n = 26)	p	d	95% low	95% high
V_{ini} (m/s)	15.77	16.08	.605	−0.14	−1.50	0.88
DTT_{ini} (m)	47.93	52.55	.348	−0.25	−14.39	5.15
DTT_{fin} (m)	6.00	4.38	.046	0.53	0.03	3.20
R^2_{time} (-)	0.93	0.92	.468	0.19	−0.01	0.03
Mean V (m/s)	10.16	10.12	.912	0.03	−0.55	0.62
SD DTT_{fin} (m)	3.38	2.35	.045	0.53	0.03	2.04
Max. dec. (m/s^2)	6.92	6.33	.174	0.36	−0.27	1.44
Onset jerk (m/s^3)	7.19	5.24	.081	0.46	−0.25	4.15

p represents the probability of observing the given means by chance if the means would actually be equal; d represents Cohen's d, a measure of the size of the difference between the means in relation to the pooled standard deviation; 95% low and 95% high represent the 95% confidence interval for the true difference of population means.

2.4 Discussion

The eight low-cost non-vestibular motion cueing systems, all of which were designed to feed back vehicle speed or accelerations, showed desired effects. During braking and cornering, the systems generally induced drivers to adopt lower speeds, exert lower vehicle accelerations (see Figure 10 for an illustration during braking), achieve higher control accuracy, and give higher subjective realism scores of the simulator. It is concluded that diverse motion cueing solutions can be successful in reducing vehicle speeds and accelerations in a fixed-base simulator.

Considering the results of individual systems, the seatbelt tensioning system caused the largest reduction in speed and deceleration and a large improvement of stopping position accuracy and precision. The seatbelt force cue is not realistic in itself; the seatbelt force in a real car does not change when decelerating during normal driving; instead it allows some freedom of movement for the driver, unless a locking mechanism is activated by sudden vehicle accelerations, in which case seatbelt forces can develop. But the seatbelt force cue as applied here provided proportional deceleration feedback with adequate intensity, and drivers were shown to be able use this feedback effectively to control the vehicle during braking.

Steering wheel and seat-based vibrations, as used to feed back longitudinal and lateral accelerations respectively, resulted in a small reduction of onset jerk during braking and a slightly lower speed and smaller SDLP through turns. It must be noted that the intensity of the vibrations was rather weak in our experiments. The steering wheel vibrations were similar to vibrations in a real car during driving on an asphalt road (as caused by road irregularities and engine vibrations; data not shown). The vibrations were presented when the driver performs the most difficult part of the task and concentrates on picking up the necessary visual and proprioceptive feedback to control the car. Therefore, in further research it is recommended to use vibrations of higher intensity to facilitate perception, inevitably making them less realistic, however.

The motion seat had a surprisingly small effect on braking and cornering performance. For braking, only the onset jerk was reduced significantly, and a reduction of stopping position error approached significance. These results can perhaps partly be explained by the fact that the seat position was linked proportionally to accelerations, and the seat movements were thus related to jerk. Apparently, these seat movements caused a stronger feedback than the actual seat angle. Perhaps another cueing algorithm for the motion seat could have resulted in more significant and desirable effects. During cornering, both the fun-ride and engineering approach resulted in slightly lower speeds. Finally, another interesting result was that the subjective realism of the simulator was increased irrespectively of what cueing strategy was used (fun-ride or engineering approach).

The screeching tire sound caused participants to avoid high accelerations during driving sessions 3 and 4 as compared to sessions 1 and 2. A disadvantage of screeching tire feedback, however, is that it is audible only during the very high accelerations the tire sound. With the sound, participants braked *later* than without the sound, and from higher speeds; they subsequently slowed down with less brake modulations, higher mean deceleration, but a smaller maximal deceleration. This indicates that drivers braked so that they searched for the limit of hearing the screeching tire sound. Because many driving simulation studies report on unrealistically fast driving, screeching sound seems a justified means to avoid the highest accelerations, but does not solve the problem of a lack of motion cues during normal driving conditions. In order to present proportional auditory feedback from low acceleration levels, the beeping sound feedback was developed. During braking, slightly lower speeds were obtained because of the beeping sound, but the deceleration profile was found to be less constant. During cornering, a lower speed through left turns, slightly smaller lane center errors, and less road excursions were found. A quasi-transfer test (Task C without feedback) showed that the group previously driving with beeps had slightly smaller SDLP values through both left and right turns and also had less road excursions. This shows that the beeps might have caused a small learning effect for steering the car smoothly through the turns.

The road noise feedback was not based on vehicle accelerations, but vehicle speed instead. The mean speed was reduced during the braking task, all other measures improved slightly as well. For the cornering task, speeds and lane center errors were also reduced, but only the decrease in lane center error through right turns reached statistical significance. When the feedback was removed in a subsequent session, the effects were reduced in magnitude, but the participants formerly driving with the speed feedback had less road departures. This shows that adding speed-dependent feedback can help to reduce vehicle speeds, as has been found before by Davis and Green (1995) and can subsequently help to reduce vehicle accelerations and increase driving accuracy and precision. This relates to the speed-accuracy trade-off; if speed is reduced the control task is less demanding, and accuracy and precision can be improved.

The pressure seat resulted in a lower standard deviation of the lateral position, and thus more constant lateral accelerations through the turn. These results suggest that participants avoided high acceleration peaks.

The use of a motion platform for driving simulation has been shown to result in changes in driver performance and an increase of subjective realism (Siegler et al., 2001). It now seems that providing acceleration feedback improves the ability to control a vehicle, irrespectively of *how* these cues are presented. In this respect, the way in which visual, vestibular, or tactile information is used by humans should be considered. The human brain uses efficient strategies to integrate and interpret visual and vestibular information (Butler, Campos, & Bülthoff, 2008; Cutting, 1997). Ernst and Bülthoff (2004) showed with a review how different sources of information are integrated by humans into the perception of acceleration and velocity. The way in which this integration takes place and the weights different sources of information are given vary substantially between individuals and experience levels. This could mean that driving performance in the simulator (vehicle speed and control accuracy) can accurately match performance in the operational environment by using "unrealistic" cues. Similarly, it could be possible that two simulators show identical driver performance, but both elicit dissimilar driver's cue weighting strategies.

The present results have important implications for researchers and practitioners who want to achieve improved simulator fidelity without using vestibular motion cues. An interesting additional thought is that alternative motion cueing systems such as those used in this paper could also be used to feed back vehicle speed and accelerations in real vehicles, to contribute to lower vehicle speed and accelerations in the real world, thereby reducing the risk of accidents.

Although the focus of this study was exclusively on in-simulator performance, one may speculate whether the results are also relevant for simulator-based training. Our quasi-transfer of training experiment with the road noise system showed that retraction of the supplementary cues had a favorable effect on the number of road departures. A potential disadvantage of using unrealistic cues during training is that the transfer to a real vehicle is disturbed, because the learner depends on cues during training which do not exist in reality. The human sensorimotor system is highly adaptive to varying sensory contributions, though. For human stance control a number of experiments have demonstrated the adaptation capabilities of the human sensorimotor control system (Nashner, Black, & Wall, 1982; Peterka & Loughlin, 2004; Van der Kooij, Jacobs, Koopman, & Van der Helm, 2001). When certain cues are missing, or others exaggerated, the human will adapt quickly to the new situation. It is therefore likely that non-vestibular motion cues in the simulator would be substituted swiftly by vestibular feedback when transferring to a real-world vehicle. Nonetheless, it remains an open question whether non-vestibular motion cueing systems result in satisfactory transfer to the real world or not (see also Groeger & Banks, 2007). This question will have to be answered empirically.

Chapter 3. The effects of route-instruction modality on driving performance in a simulator

Abstract

Currently, most instructions given during driver training in simulators are presented in the auditory modality through speech. This may be done to simulate a human instructor and because the auditory instructions interfere little with the predominantly visual driving task. However, simulators provide the possibility to present instructions to the drivers in various other modalities. This research aims to give insight into how driving performance is influenced by presenting route instructions in the visual rather than the auditory modality.

An experiment was performed on The Dutch Driving Simulator in which the route-instruction modality was altered between the auditory modality (speech), the visual modality (projected arrows), and a combination of both. During the experiment, beginner-level instructions concerning the driving and traffic situations were presented using speech. The subjective experience of the participants was measured using questionnaires.

The visual and multimodal route instructions resulted in significantly fewer turn errors and turn indicator errors compared to the auditory route instructions. The multimodal route instructions resulted in the fastest response times, an effect which was most pronounced for participants who drove faster and for participants who had (or reported they had) poorer driving skills. Most participants preferred the visual route instructions. This research shows that although visual instructions interfere with the predominantly visual driving task, under certain conditions they can result in better driving performance.

De Groot, S., De Winter, J.C.F., Mulder, M., Kuipers, J., Mulder, J.A., &Wieringa, P.A. (2006). The effects of route-instruction modality on driving performance in a simulator. *Paper presented at the 9th TRAIL congress,* Rotterdam, The Netherlands.

3.1 Introduction

The use of driving simulators for driver training provides several potential advantages compared to driver training on the road (Kappé & Emmerik, 2005). Safety, objective measurements, environments that are purposefully built for learning, and cost savings are the most important advantages. Simulators also provide the possibility to present instructions and provide feedback to the drivers in various modalities. Concerning driver training, this is not (yet) possible in real cars, where the human driving instructor mainly uses speech to transfer the required information to the students.

Currently, most instructions given by the Dutch Driving Simulator (Green Dino, 2008) are presented in the auditory modality through speech. This is done to simulate a human instructor (Weevers, Kuipers, Brugman, Zwiers, Van Dijk, & Nijholt, 2003). In addition, the spoken instructions are believed to interfere little with the predominantly visual driving task (cf. Wickens, 1999). However, a disadvantage of auditory information transfer through speech is that only one instruction can be presented at a time, restricting the information transfer rate (Reed, 1998). Because a lot of information may have to be presented to learner drivers, the question arises whether it is beneficial to present some of the information in alternative modalities.

Visual displays can present multiple messages simultaneously and messages can be presented for longer periods of time, so that the drivers can use the information when they are ready to perceive and process it. A visual display thus allows for self-paced information processing. However, visual displays interfere with the predominantly visual driving task. In addition to the option of presenting information in the auditory or visual modality, information can also be presented in multiple modalities simultaneously. In general, it can be expected that responses to redundantly presented information are improved compared to information presented in a single modality. However, for complex tasks such as car driving this general principle might not always be valid (cf. Van Esch, 2001). Whether and to what extent presenting redundant information can be used during simulation based car driving is an important question for the development of driving simulators.

This research aims to give insight into how driving performance is influenced by relieving the auditory modality from specific verbal instructions and presenting them in an alternative modality. An experiment was performed wherein the route instructions were presented in either the auditory modality, the visual modality, or a combination of both. During the experiment the remaining beginner level instructions were presented using speech as usual on the Dutch Driving Simulator. Experienced drivers were used for the experiment, to minimise learning effects and focus on the influence of display modality on driving performance.

3.2 Method

3.2.1 Apparatus

The experiment was conducted on the Dutch Driving Simulator, which is a low-cost fixed-base driving simulator (www.rijsimulatie.nl). The simulator was operated using all controls associated with normal car control: steering wheel, accelerator, brake, clutch, gear shift lever, and turn indicator lever. Force feedback was provided on the steering wheel to represent the self-aligning torque of the front wheels. The front view projection had a resolution of 1,024 x 786 pixels; the side-view projections had resolutions of 800 x 600 pixels. An extra projector was used to present the visual route instructions in front of the driver. The dashboard (which included engine rpm, speed and a turn indicator light), the vehicle interior and mirrors were integrated in the projected image. The auditory information was presented non-directionally. The driver control signals and vehicle states were logged with a frequency of 50 Hz.

3.2.2 Participants and instructions to participants

Ten participants (five females, five males) between the age of twenty-one and twenty-four (mean age 22.8) who all possessed a driving license for a minimum of two years joined the experiment. Experienced participants were used, because it was hypothesized that the effects on driving performance of reducing the amount of spoken instructions would be too large for beginner drivers. Paradoxically, beginner drivers may need more instructions than experienced drivers because they still have much to learn, but because they are mainly concentrated on basic car control they have more difficulties processing these instructions (Groeger & Clegg, 2004). The participants were asked to complete the route using their normal driving style, without committing traffic violations. They were asked to always use the turn indicators after the route instructions. However, it was not emphasized to do this immediately.

After the experiment, the participants were asked to complete a questionnaire concerning the experiment and a questionnaire concerning the Index of Learning Styles (ILS) (Soloman & Felder, 2005; Felder & Spurlin, 2005). The ILS was included to gain insight into the participants' preferred input modality for information. This experiment was not intended to be a learning task, but the preferred input modality in the ILS may be correlated with the preferred instruction modality during driving.

During the experiment a virtual instructor presented beginner-level instructions in the auditory modality through speech. We choose for the beginner-level, because this means the simulator presents many instructions and comments concerning the driving performance and the traffic situations that were encountered (e.g., intersections and curves).

3.2.3 Experiment procedure

After a learning period of five minutes, included to allow the participants to adapt to driving in the simulator (McGehee, Lee, Rizzo, & Bateman, 2001), the measurement sessions were started. Each participant drove three sessions of 10 minutes each. For each session the modality for the route instructions was changed between three experimental conditions in randomized order (see paragraph 3.2.4). Participants drove in an urban environment containing many right priority intersections without traffic symbols or lights. At the intersections the participants were randomly instructed to turn left, right, or go straight. During the entire experiment a speed limit of 50 kilometers per hour was instructed, according to the Dutch speed limit for urban areas. To improve realism, some other traffic drove around as well. Participants were allowed to overtake slow driving traffic.

3.2.4 Independent variables and display design

The independent variable was the display type. Three types of display were tested: auditory route instructions were presented as speech during condition 1 (C1), visual route instructions were presented as arrows during condition 2 (C2), and a combination of speech and visual route instructions that are presented simultaneously during condition 3 (C3). The auditory route instructions (C1) consisted

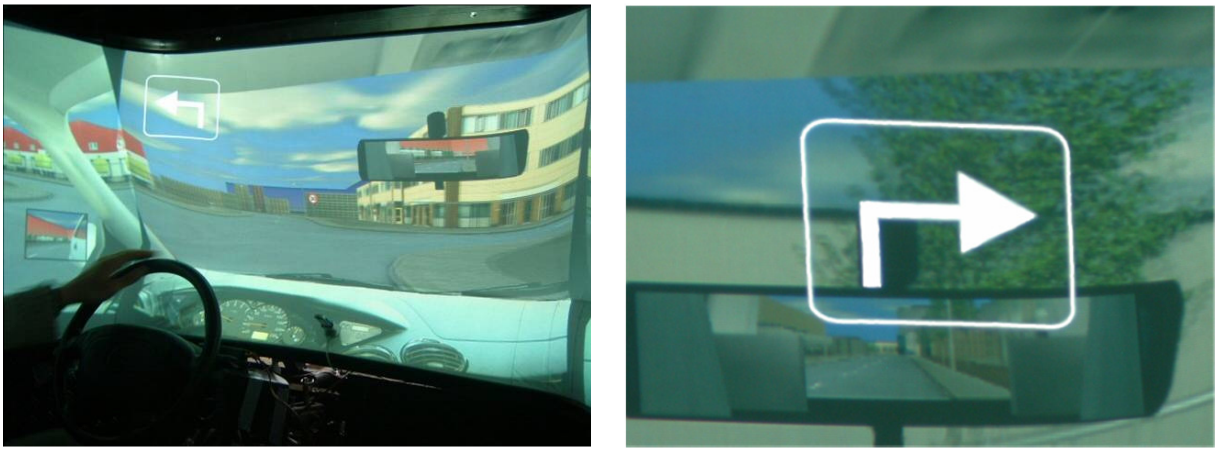

Figure 1. Visual route-instruction display (C2 and C3). The left picture shows the display for left turns, the right picture shows the display for right turns.

of a human pre-recorded voice instructing the participants to turn left or right at the next intersection. The visual route instructions (C2) consisted of an arrow pointing left or right. The left pointing arrow was projected in the upper left corner of the front projection, the right pointing arrow in the upper middle part of the front projection, near the rear view mirror. The projected arrow was positioned so that the participants did not need to look directly at the symbol and could retain the visual focus on the road. Each arrow was presented for 6 seconds. Figure 1a presents a view of what the driver saw at an intersection where a left turn had to be made. Figure 1b shows the right-turn symbol just above the rear-view mirror.

The combination display (C3) was constructed by combining the route instructions of C1 and C2. Concerning the timing of the projected arrows during C2 and C3, the route-instruction arrows were displayed at the first moment that a left/right distinction could be extracted from the auditory route instruction.

3.2.5 Dependent measures

Table 1 shows the dependent measures. The dependent measures were divided into speed and accuracy measures (cf. De Winter, et al., 2006; Zhai, Accot, Woltjer, 2004). The turn indicator response time was included as a measure related to the time needed to process the route instructions. For the measures Speed, SDLP, and SDSteer a second measure was calculated for the time interval between the route instruction and the start of the nearest intersection (these second measures are defined as call measures). Subjective measures were obtained using the ILS questionnaire (Soloman, 2005) and the questions presented by Table 2.

3.2.6 Statistical analyses

The dependent measures were tested for significant between-condition differences using a repeated-measures Analysis of Variance (ANOVA). Because a normal distribution and equal variance cannot be assumed for both the IndicatorErrors and TurnErrors measures, the non-parametric Friedman test was used for testing the between-condition of these measures. A significance level of .05 was used for all tests. A Spearman intercorrelation matrix was constructed of all objective and subjective measures for all conditions. The correlations of 0.65 and higher ($p < .05$) were selected for further analysis.

Table 1. Dependent measures which were calculated for each session.

Speed	
Speed [km/h]	Mean speed
Speed_call* [km/h]	Mean speed on call-intervals
Calls [#]	The number of route instructions
Tcross [s]	Mean of the times taken to cross the intersections
Accuracy	
SDLP [m]	Standard deviation of the lateral road position
SDLP call* [m]	Standard deviation of the lateral road position on call-intervals
SDSteer [deg]	Standard deviation of the steering wheel angle
SDSteer call* [deg]	Standard deviation of the steering wheel angle on call-intervals
IndicatorErrors [%]	Percentage of calls that the indicator was wrongly used
TurnErrors [%]	Percentage of calls that the route instruction was not obeyed
Indicator response	
ResponseTime [s]	Time from route instruction to turn indicator execution

*A call-interval was defined as the interval from the presentation of the route instruction to the moment of driving onto the intersection. The session values of the call measures were calculated both as the mean and median of the calls per session.

Table 2. Subjective measures.

Q1 [%]*	Did you enjoy driving in the simulator? (anchors: not at all, very much)
Q2 [%]*	Do you make errors between left and right in everyday life? (anchors: never, all the time)
Q3 [%]*	Did it take you long to adapt to driving in the simulator? (anchors: very long, very short)
Q4 [%]*	Was the 5 min adaptation time long enough for you? (anchors: far too little, far too much)
Q5 [%]*	How was your concentration level during the experiment? (anchors: bad, good)
Q6 [%]*	How was the steering wheel control in the simulator? (anchors: very hard, very easy)
Q7 [%]*	How well could you estimate where you needed to brake before the next turn or intersection? (anchors: very bad, very well)
Q8 [%]*	Did you remember to turn on the turn indicators quickly after the route instructions? (anchors: never, always)
Q9 [%]*	How did you experience the spoken route instructions? (anchors: very vague, very clear)
Q10 [%]*	How did you experience the route instructions via the projected arrows? (anchors: very vague, very clear)
Q11 [%]*	How did you experience the route instructions via the combination of speech and projected arrows? (anchors: very vague, very clear)
Q12	Which of the route instructions did you prefer: C1: Only speech C2: Only projected arrows C3: Combination of speech and projected arrows

*Questions had to be graded between 1 and 10, corresponding to the anchors. This grade was transformed to a percentage, where 1 corresponds to 0% and 10 corresponds to 100%.

3.3 Results

3.3.1 Condition-means

Table 3 presents the means of all objective measures for the three conditions.

Table 3 reveals that there are significant between-condition differences for the Mean ResponseTime and the IndicatorError measures. A Tukey-Kramer test indicated that C1 resulted in significantly more indicator errors than C2 or C3. The Mean ResponseTime was lower during C3 than C1 or C2. Figure 2a presents the Mean ResponseTime among participants for the three conditions. The dash-dotted lines were plotted to illustrate the effects for each participant.

Figure 2b presents the IndicatorErrors for all participants for the three conditions. It can be seen that most indicator errors occurred during C1 and that some participants accounted for many more indicator errors than others. Figure 2c presents the Speed for all participants for the three conditions. No significant differences existed between the conditions, but there were large individual differences. Figure 2d presents Median SDSteer call for all participants for the three conditions. As with the Speed measure, Median SDSteer call shows large individual differences and relatively small differences between the three conditions.

Table 4 presents the means among participants of the questionnaire completed after the experiment. On average, participants reported to adapt quickly to driving in the simulator (Q3) but could have done with a longer learning period (Q4). They reported good concentration levels (Q5) and felt like they were moderately in control of steering and braking (Q6, Q7). For this experiment, queries Q9 to Q12

Table 3. Condition-means for the three conditions.

	C1		C2		C3		Two-way ANOVA	
Speed	*M*	*SD*	*M*	*SD*	*M*	*SD*	*p*	
							Participant	Condition
Speed [km/h]	36.6	6.7	35.6	4.8	36.8	6.4	<.001	n.s.
Mean Speed call [km/h]	38.0	9.1	36.6	5.9	37.6	6.8	<.001	n.s.
Median Speed call [km/h]	39.0	9.0	37.6	6.2	39.0	7.1	<.001	n.s.
Calls [-]	17.8	2.8	17.7	2.5	17.8	2.5	.001	n.s.
Mean Tcross [s]	3.66	0.68	3.66	0.30	3.86	1.21	.051	n.s.
Median Tcross [s]	3.42	0.50	3.52	0.38	3.42	0.59	.008	n.s.
Accuracy								
SDLP [m]	0.60	0.08	0.57	0.09	0.59	0.10	.003	n.s.
Mean SDLP call [m]	0.60	0.13	0.54	0.11	0.57	0.10	.010	n.s.
Median SDLP call [m]	0.54	0.10	0.51	0.10	0.51	0.10	.015	n.s.
SDSteer [deg]	62.9	8.6	61.4	2.5	62.7	5.5	.030	n.s.
Mean SDSteer call [deg]	24.4	3.3	22.5	3.7	22.4	3.7	.017	n.s.
Median SDSteer call [deg]	22.2	2.6	22.1	3.3	21.0	3.4	<.001	n.s.
IndicatorErrors* [%]	5.8 2A,6B 1C,1D	7.3	0.6 1D	1.9	1.1 1A,1B	3.3		.022**/***
TurnErrors [%]	0.59	1.86	0.59	1.86	0	0		n.s.***
Indicator response								
Mean ResponseTime [s]	2.64	0.63	2.49	0.51	2.04	0.39	n.s.	.027**
Median ResponseTime [s]	2.11	0.39	2.15	0.26	1.87	0.31	0.015	.043**

*Four different kind of indicator errors occurred after the route instructions:
A: first correct, then wrong.
B: first wrong, then correct.
C: no indicator used at all.
D: no indicator used at all, but also no turn (counted also as turn error).
**To test which condition-mean is significantly different to the others, a Tukey-Kramer pairwise comparison test was performed.
***The Friedman test was used for testing the between-condition effects.

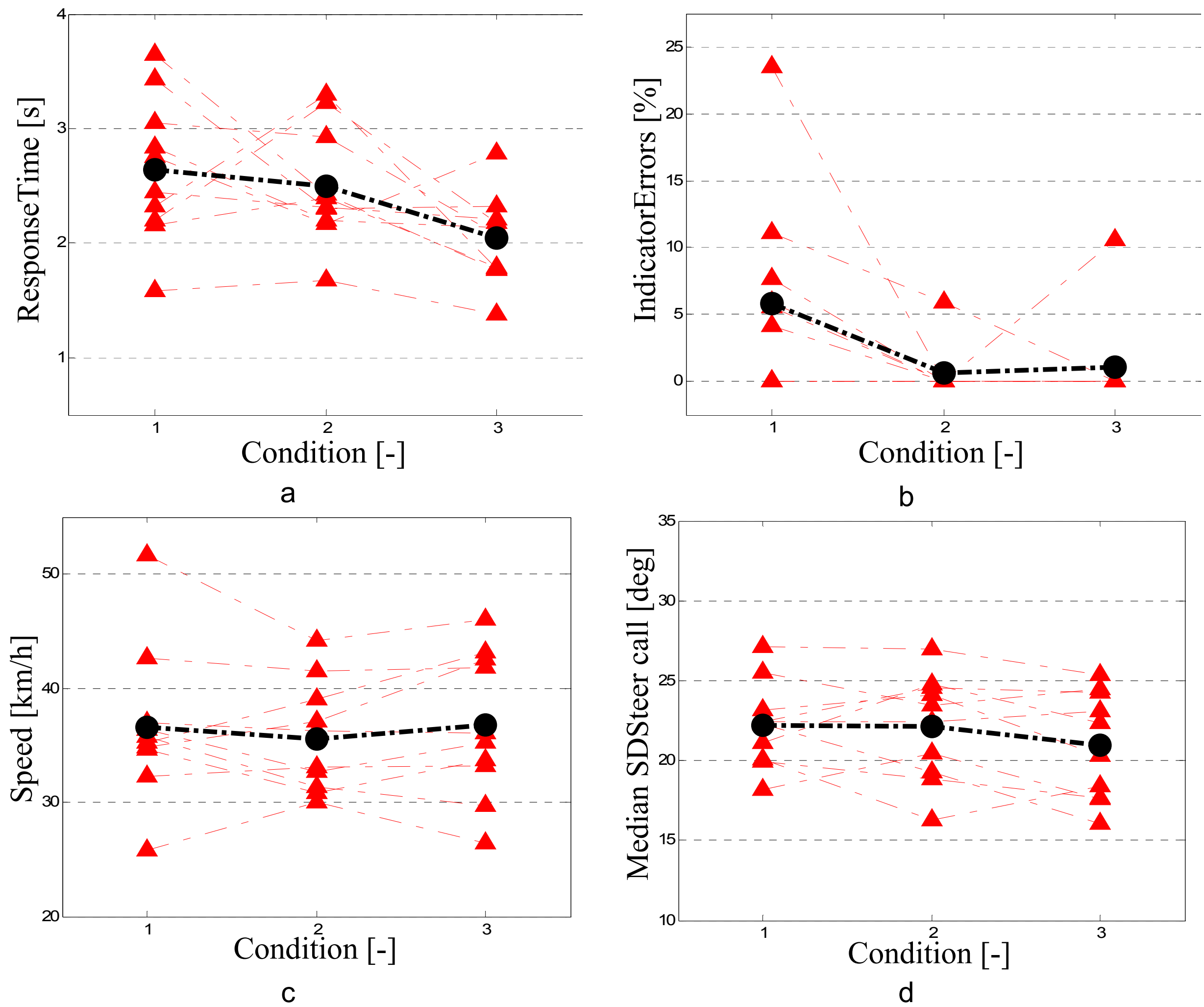

Figure 2. (a) ResponseTime, (b) IndicatorErrors, (c) Speed, and (d) Median SDSteer call for the three conditions. The triangles show the individual participant values for all conditions, the circles represent the condition-means. Condition 1 = speech, Condition 2 = projected arrows, Condition 3 = combination of speech and projected arrows.

were of particular interest. These questions revealed that C3 was experienced by the participants as being most clear, while 6 out of the 10 participants indicated that they preferred C2.

3.3.2 Subjective experience correlation with objective data

By combining the subjective and objective measures to construct a Spearman intercorrelation matrix, a number of interesting explorative results were obtained. Participants who experienced C1 as being clear (Q9) had significantly lower Speed, Speed call, and fewer Calls during C1. Figure 3a shows the relationship between the mean Speed call and the response to Q9. Participants who experienced C2 as being clear (Q10) had significantly higher Mean ResponseTime during C1, fewer IndicatorErrors, and fewer TurnErrors during C2. These participants may have had difficulties with processing the auditory route information. Figure 3b shows the Mean ResponseTime during C1 versus the response to Q10.

Participants who experienced the route instructions with C3 as being clear (Q11) had significantly higher SDLP and SDLP call values when driving during C3.
Previous research has shown that high SDLP values may be indicative of poor vehicle handling abilities (De Winter et al., 2006) or high workload (Summala, 1996).

Table 4. Means (*M*) and standard deviations (*SD*) of questionnaire responses. C1 = speech, C2 = projected arrows, C3 = combination of speech and projected arrows.

Question	*M*	*SD*	Question	*M*	*SD*
Q1 [%]	70	32	Q7 [%]	47	27
Q2 [%]	29	38	Q8 [%]	87	10
Q3 [%]	69	19	Q9 [%]	76	25
Q4 [%]	48	12	Q10 [%]	79	26
Q5 [%]	77	18	Q11 [%]	90	16
Q6 [%]	56	21	Q12 [C1, C2, C3]	1xC1, 6xC2, 3xC3	

In other words, the participants who may have poor vehicle handling abilities or high workload experienced the multimodal display (C3) as being clear.

The observed SDLP scores are high (about 0.6 m) compared to SDLP values found in literature for a normal driving task (typically about 0.3 m). This can be explained by the nature of the driving task, which involved a lot of curved road segments and some slow driving traffic that could be passed by the participants. In other experiments on the Dutch Driving Simulator lower values for SDLP have been obtained for driving on straight road segments.

Participants who indicated that they often make left/right errors in every-day life (Q2) had a high mean SDLP call during C1 and C2, but not during C3. Q2 also correlated with TurnErrors during C1, indicating that participants who reported more left/right turning errors in everyday life also made more left/right turning errors in the simulator. Figure 4 illustrates the correlations of mean SDLP call versus Q2. The ILS input score (visual/verbal) was related to the Mean ResponseTime during C2. Participants who had a higher verbal score had a higher Mean ResponseTime when driving with the visual-only instructions. The relationship between the Mean ResponseTime during C2 and the ILS input score is presented in Figure 5.

3.4 Discussion and conclusions

Although this study was of limited scope and was performed with only 10 participants, it indicated the potential of using alternative modalities during simulator-based driver training.

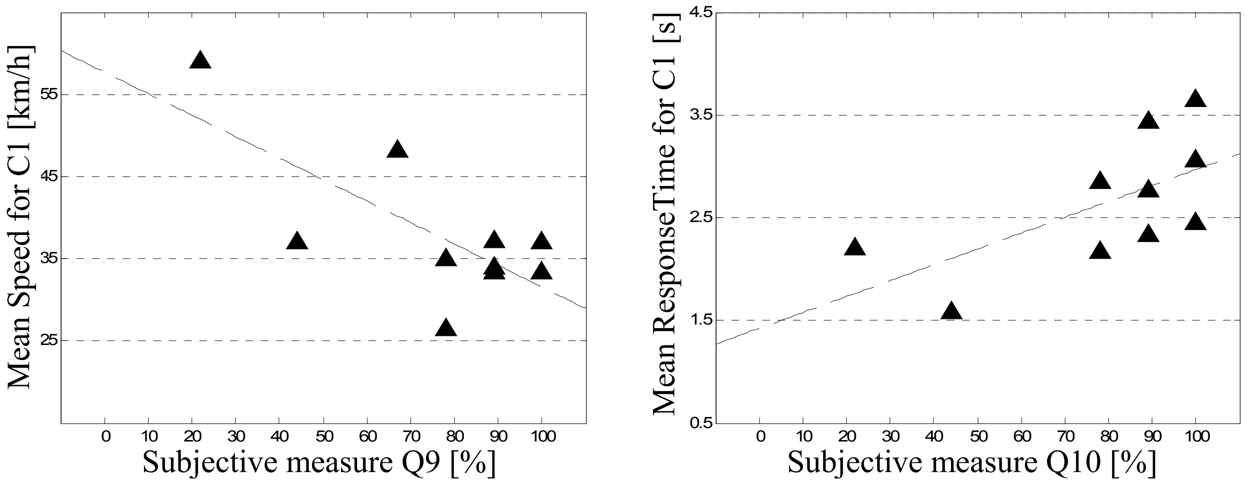

Figure 3. Mean Speed call for C1 versus Q9 for all participants on the left and Mean ResponseTime for C1 versus Q10 for all participants on the right. The triangles show the participant‘s individual values, the dashed line represents a linear fit. C1 = speech.

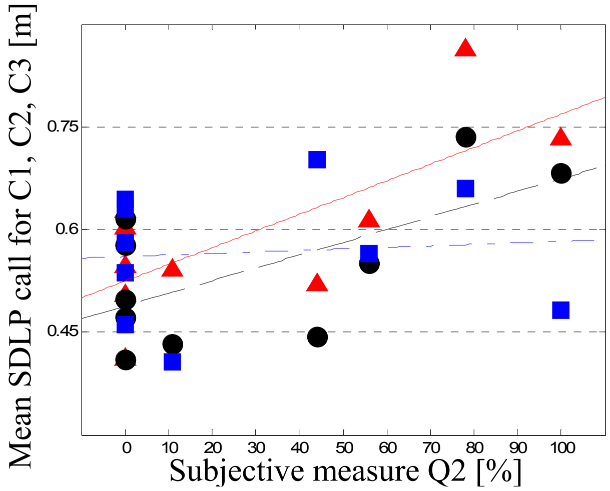

Figure 4. Mean SDLP call during C1, C2, and C3 versus Q2. The triangles and solid line represent C1, the circles and dashed line represent C2, the squares and dash dotted line represent C3. The lines represent linear fits. C1 = speech, C2 = projected arrows, C3 = combination of speech and projected arrows.

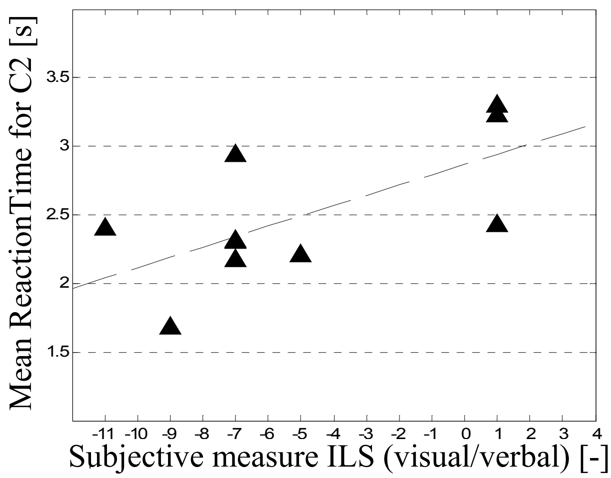

Figure 5. Mean ResponseTime during C2 (projected arrows) versus ILS score on the visual-verbal dimension. The dashed line represents a linear fit.

3.4.1 Differences between displays

The combination of visual and auditory route instructions (C3) resulted in lower indicator response times compared to visual-only (C2) or speech-only (C1) route instructions. The lower indicator response times during C3 are likely to be caused by a redundancy gain, as literature indicates (Wickens, 1999; Van Esch, 2001).

Both C2 and C3 reduced indicator errors compared to C1. The reduced indicator errors for C2 and C3 may be caused by the large period (6 seconds) in which the visual information was available to the drivers during C2 and C3. Whenever the driver felt ready to execute the indicator, this could be done confidently. This is in contrast to the auditory information during C1, of which the uncertainty about the information grows with increasing time after the instruction.

No differences in driving speed or accuracy were found between the three display designs. This suggests that the mental workload was not affected by the different route instruction displays, in agreement with Summala's model of behavioral adaptation (Summala, 1996).

3.4.2 Differences between participants

The speed and accuracy varied considerably among participants. When the driving speed is higher, more events are encountered per time unit and more instructions are presented by the simulator. The explorative analysis of the correlation matrix indicated that the higher the speeds of the participants, the more unclear the auditory route instructions during C1 (speech only) were experienced. The faster driving participants experienced more events, consequently received more instructions, and may have therefore had more difficulty with extracting the route information from all the auditory instructions.

Concerning accuracy, the participants who were less accurate found the route instructions of C3 more clear. Based on earlier findings less accurate participants are assumed to have poorer vehicle handling abilities (De Winter et al., 2006) or to have a higher subjective workload (Summala, 1996). Also, the results indicated that participants who reported to be bad in everyday left/right distinctions were most helped by the redundant auditory and visual information during C3. It can be reasoned that participants with less-developed driving skills are more involved with basic car control and have less spare capacity to absorb route instructions. The redundant information then helps to focus attention on the information that is presented.

It is remarkable that the clearest and most effective route instruction modality, the combination display of C3, was not the most preferred one. Instead, the participants preferred the visual display of C2. A possible explanation is that because C2 did not provide additional auditory route instructions, there was some relief from the auditory modality.

3.4.3 Multiple resource theory

The multiple resource theory (Wickens, 1999) can be of help to analyze the task interferences during simulation-based driving. The theory distinguishes between processing stages (perceptual vs. response execution), perceptual modalities (auditory vs. visual), and processing codes (verbal/digital vs. spatial). The primary driving task (vehicle control) demands visual-spatial resources for perception and processing.

The multiple resource theory predicts that a secondary task demanding auditory/verbal or visual/verbal resources will interfere less with the driving task than those that compete with the visual/spatial resources demanded by the primary driving task. The Dutch Driving Simulator currently presents a lot of auditory/verbal information per time unit at beginner instruction level. Consequently, any information that can be presented as visual/verbal information relieves the auditory modality, which reduces errors for the information presented visually/verbally and possibly also for the remaining auditory information.

3.4.4 Implications for simulation-based driver training

Using alternative modalities for instructions shows potential to increase learning efficiency during simulation-based driver training. Based on the results of the present study, the following implications for simulation-based driver training can be formulated: Although the visual route instructions may have interfered with the predominantly visual driving task, they resulted in less turn errors. This might be an indication that too many speech instructions are presented or that the visual route instructions have other advantages (such as self-pacing).

When the attention of students needs to be focused on instructions that are more important to the learning task than other instructions, these instructions can be presented in two or more modalities simultaneously. Because simulators provide the possibility of presenting instructions to the drivers in various modalities, and large interpersonal differences between driver students exist, it seems attractive to adapt the instruction modality to the individual characteristics of the students.

Visual displays have the potential to improve the subjective acceptance of training in a simulator, which in turn is likely to lead to more efficient training. The next step will be to study the effect of changing the instruction modality on a learning task. If the auditory modality is relieved from recurrent speech instructions, it is hypothesized that the auditory feedback that remains receives more attention from the students which in turn increases training effectiveness.

Part 2

Simulator-based driver training

Chapter 4. Didactics in simulator-based driver training: current state of affairs and future potential

Abstract

Moderate-fidelity driving simulators are increasingly being used for cost-effective initial driver training. Apart from the need to satisfy simulator fidelity requirements, more attention is needed on the didactical properties of the training programs in order to yield more effective training. This paper investigates the didactical properties of current driver training simulators, and provides recommendations for improving the instructional design. A survey shows that the intelligent tutoring systems of current driver training simulators are mostly imitating the human instructor and that the "first principles of instruction" (Merrill, 2002a) are not implemented to their full potential. Hence, there is ample room for improvement of the didactical properties by fully exploiting the many visualization, demonstration, and performance-assessment opportunities provided by modern driving simulators. Furthermore, objective performance ratings of students can be used to provide accurate and consistent feedback-on-performance, something that is not possible in real cars, but which is often essential for effective skills training. It is recommended to use empirical experimentations to improve the instructional design of simulator-based driver training for specific learning outcomes and validate the use of the first principles of instruction to facilitate learning.

De Groot, S., De Winter, J. C. F., Mulder, M., & Wieringa, P. A. (2007). Didactics in simulator-based driver training: current state of affairs and future potential. *Proceedings of the Driving Simulation Conference North America*, Iowa City, IA.

4.1 Introduction

An increasing number of driving schools are integrating simulators in their driver training curricula. The potential of simulators for driver training may be improved if more emphasis is placed on the didactical aspects of the training program during development.

Driving simulators are developed for various purposes, such as research, entertainment, or training. As far as training is concerned, simulators provide several advantages as compared to driving on the road. Examples are the measurement of driving performance (Hoeschen et al., 2001), the possibility to train in purpose-developed virtual environments (Hoeschen et al., 2001), and the increased opportunities to present information and provide feedback using various modalities. Automated training offers further economic advantages, if the human driving instructor is replaced by an automated tutoring system. However, in situations where a human teacher may adapt and compensate for flaws in the learning materials, the learning materials have to be of excellent quality if they are to stand on their own (Barnard, 2006). Several researchers have indicated that the didactical aspect of simulator-based training programs should receive more attention (e.g., Kappé & Emmerik, 2005; Salas, Bowers, & Rhodenizer, 1998). Examples of questions that need to be addressed regarding the didactical aspects are: “What are the didactical properties of simulator-based driver training at this moment?”, and “How can the didactical properties of simulator-based driver training be improved?” To answer these questions, a way to describe the didactical properties is necessary.

4.1.1 Didactics in driver training

The effectiveness of a training program can be determined quantitatively (how much is learned given a certain amount of training time) or qualitatively, by investigating, for example, the instructional design of the program. A quantitative analysis was performed with data from one specific driving school which used an additional human instructor to assist the automated training on their simulators. Data from this school was compared with data from other driving schools using the same simulator. Results are presented in Table 1. The scores were normalized so that the mean is equal to zero, and the standard deviation is equal to one for the complete group (N = 859).

Results show that more tasks are executed erroneously, fewer tasks are executed correctly, and, on average, tasks take more time to complete when the human instructor is present. This analysis shows that students training on systems using an intelligent tutoring system do not necessarily benefit from extra human attention. Instead of helping as he intended, the human instructor seems to have delayed the training process with extra instructions and feedback, and introduced extra information to process for the students. Although there are some methodological weaknesses, such as lack of experimental control and possible self-selection bias, this case does suggest that automated training may outperform training by a human instructor and that simulators have potential to standardize and objectify training programs. Additionally automated training programs can benefit from using proven educational techniques which are used methodologically to develop them.

4.1.2 Goal of this article

This article will qualitatively analyze the quality of instructional designs of driver training simulators and provide suggestions for improvement. The evaluation includes the instructional designs of four selected moderate-fidelity driver training simulators

Table 1. Differences between automated instruction with and without human help.

	Driving school with human help (n=39)		Other driving schools (n=820)		
	M	*SD*	*M*	*SD*	*p*
Error-score*/** [-]	0.05	0.49	0.00	1.02	.772
Error percentage-score*/*** [-]	0.31	0.57	−0.01	1.01	.047
Correct-score*/**** [-]	−1.39	0.72	0.07	0.96	<.001
Speed-score*/***** [-]	−1.70	0.83	0.08	0.93	<.001
First practical exam pass rate [%]	49	51	53	50	.618
First sim. lesson to license [days]	235	86	190	96	.006

*Measures obtained from De Winter et al. (2007).
**The error-score is based on number of erroneously executed tasks.
***The error percentage-score is based on the number of errors with respect to the total number of attempts.
****The correct-score is based on the number of correctly executed tasks.
*****The speed-score is based on the average task time.

which are available for Dutch driving schools: 1) The ANWB simulator (including the supporting computer-based training lessons) (www.anwb.nl) developed by ST-software (www.stsoftware.nl), 2) The DriveZone simulator (DZ) (www.drivezone.nl) developed by VSTEP (www.vstep.nl), 3) The Dutch Driving Simulator (DDS) developed by Green Dino (www.greendino.nl), and 4) The VR-systems simulator (VR) (www.vrsystems.nl) developed by dr. Foerst (www.drfoerst.de). The article is intended to give a general impression of the current state of affairs concerning the didactics of commercially available driver training simulators, not to compare or classify the evaluated simulators.

4.2 Method

The instructional design of the simulator-based driver training programs will be analyzed using the “first principles of instruction” (Merrill, 2002a). Merrill derived the first principles from the major instructional design theories and models and are applicable for both knowledge as well as (motor-) skill acquisition. The first principles are presented here as questions, just as in Merrill (2002b), because this allows for an easier analysis. If there is an affirmative answer to the questions, learning is incrementally improved. The first principles of instruction (Merrill, 2002b):

1. Problem-centered: Is the courseware presented in the context of real-world problems? Are learners shown the problem, engaged at the task as well as the operation level, and involved in a progression of problems?
2. Activation: Does the courseware attempt to activate relevant prior knowledge or experience? Are learners directed to recall relevant past experience or provided relevant experience? Are they encouraged to use some organizing structure?
3. Demonstration: Does the courseware demonstrate what is to be learned rather than merely giving information about what is to be learned? Are the demonstrations consistent with the instructional goals? Is learner guidance employed? Do media enhance learning?
4. Application: Do learners have an opportunity to apply their newly acquired knowledge or skill? Is the application consistent with the instructional goals, and does it involve a varied sequence of problems with feedback? Are learners provided with gradually diminished coaching?

5. Integration: Does the courseware provide techniques that encourage learners to integrate (transfer) the new knowledge or skill into their everyday life? Do learners have an opportunity to publicly demonstrate their new knowledge, reflect on their new knowledge, and create new ways to use their new knowledge?

The first principles of instruction in an instructional design can be used as a framework to analyze the didactical properties of the design. The discussion in the next section aims to answer the following two questions, along the lines of the first principles of instruction:

- What are the didactical properties of current instructional designs?
- What are the possibilities to improve the instructional designs?

4.3 Results

Principle number one, problem-centered

Learning is improved when learners are engaged in solving real-world problems. Merrill uses the word ‘problem’ to indicate a wide range of activities, with the most critical characteristics being that the activity addresses the whole task rather than only components of this task, and that the task is representative for those tasks which the learner will encounter in real life.

<u>Is the courseware presented in the context of real-world problems?</u>

Currently: Most simulators start their lessons with a spoken and/or written message containing the general learning goal of the exercise, sometimes making use of pictures of real world situations (ANWB), followed by procedural comments on the required actions of the students during the lesson, not with a clear presentation of a real-world problem (ANWB, DZ, DDS).

Possibility: Students should be shown the whole task as it is performed in the real world to solve a relevant problem. A movie could be shown, including verbal and/or visual supporting information.

<u>Are learners shown the problem, engaged at the task as well as the operation level, and involved in a progression of problems?</u>

Currently: Simulator-based driver training programs focus on the operation level, and do not explain why certain procedures are valuable, or what happens if the procedures are not followed correctly. For example, during the drive-away lesson on the DDS, the instructions explain when the clutch should be operated and how fast, without informing the student why it needs to be done like that or what happens if the clutch is operated differently.

Possibility: The students should be involved at the problem level, not just at the operation, procedure or action level. This can be done by explaining the whole problem and associated difficulties. For the drive-away task described above (at the current situation), the function of the clutch must be explained, including possible results of too fast or too slow clutch pedal release.

Principle number two, activation

Learning is facilitated when existing knowledge is activated as a foundation for new knowledge or skills.

<u>Does the courseware attempt to activate relevant prior knowledge or experience?</u>

Currently: The introduction of the lesson is used to explain students the goal of the lesson, but it generally fails to check whether the required skills are present or to activate skills required for the exercise (ANWB, DDS, VR).

Possibility: An exercise could be performed prior to the actual lesson to test some of the required skills to perform the lesson. Results of this test could be used to adjust the starting level difficulty of the lesson to the individual skills of the student, and activates the skills and knowledge necessary for the lesson.

Are learners directed to recall relevant past experience or provided relevant experience?

Currently/possibility: When there is no possibility to use relevant past experience as a base for new knowledge or skills, relevant experience must be provided. The gear-change lesson on (DZ) is a good example of this kind of activation. Students performing this exercise are often real beginners, without knowledge of the car controls. The gear-change lesson starts with part-task training to familiarize them with the operation of the gear lever, using a simple screen showing the H-pattern of the shifting lever. Before gear changes are practiced including lateral control of the vehicle, familiarization with the relevant individual controls is established.

Principle number three, demonstration

Learning is promoted when the instruction demonstrates what is to be learned, rather than merely 'telling' the student what is to be learned.

Does the courseware demonstrate what is to be learned rather than merely telling information about what is to be learned?

Currently: Demonstrations are used in many instructions, but there is no structural integration in all of the exercises. Also, demonstrations are often slow-motion action sequences, focusing on the task procedure. Although the procedure or action sequence itself becomes clear during the demonstration, it neither gives the student a clear idea of what a successful task completion looks like, nor what the consequence of bad execution is (DDS).

Possibility: Demonstrations can be used in all lessons to give the students a clear idea of what is required from them. Next to the slow-motion demonstrations to explain the correct procedures, also real-time demonstrations showing reference or goal behavior could be useful.

Are the demonstrations consistent with the instructional goals?

Demonstrations must be consistent with the intended learning outcome. Gagné (2005) defined five learning outcomes: 1) intellectual skill, 2) cognitive strategy, 3) verbal information, 4) attitude, 5) motor skill; each requiring different learning conditions for effective learning. Both Gagné (2005) and Merrill (1997) identified how to adapt training to specific learning outcomes.

Currently/possibility: Procedure training, as defined by Merrill (1997), requires a demonstration which presents the student with the task to be accomplished, with a list of the steps and their order. Each step must be demonstrated in detail. If the procedure is applicable in various situations, demonstrations of additional cases are required. The drive-away task is a good case where procedure training is relevant. The Dutch Driving Simulator demonstration during the drive-away lesson contains most elements needed for the procedure learning outcome, focusing on the order of execution and highlighting the important aspects to the students.

Is learner guidance employed?

Currently/possibility: During the demonstration the learner needs guidance to direct attention towards structural features of the task. During the demonstrations that are presented on simulators, attention is generally focused by zooming in, using slow-motion videos of action sequences or showing detailed pictures (ANWB, DZ, DDS).

Do media enhance learning?

Currently/Possibility: The modalities in which information and feedback are presented should be tuned for optimal information transfer. During most instructions visual images or movies are combined with a spoken voice. Sometimes this is extended with redundant visual text messages. Presenting the information both visually and verbally is preferred, as different people have different optimal information input channels.

Principle number four, application

Learning is promoted when learners are required to use their new knowledge or skill to solve problems (practice).

Do learners have an opportunity to apply their newly acquired knowledge or skill?

Currently: Simulators are mainly used to practice the complete driving task, just like it is practiced in the real world (ANWB, DZ, DDS, VR).

Possibility: More focus on the application of new knowledge or skills is possible during the application phase. The elements of focus in a specific lesson deserve extra attention concerning the feedback and general exercise design.

Is the application consistent with the instructional goals, and does it involve a varied sequence of problems with feedback?

Currently: Varying types of environments are used for different instructional goals. Variation during the lesson is often implemented through the behavior of other vehicles which show up at different locations on the virtual road. Feedback is generally limited to extrinsic feedback in the form of auditory verbal feedback from a (virtual) instructor (ANWB, DZ, DDS) or text messages appearing on the screen (VR).

Possibility: The practice must be consistent with the intended learning outcome. Concerning procedure training, Merrill (1997) states that both intrinsic feedback, observing the consequences of a given action or set of actions, and extrinsic feedback, informing the student about the appropriateness of a given action or set of actions, should be available. Especially augmentation of the intrinsic feedback (for example highlighting lines or displaying performance scores after tasks) is a potential advantage of simulator-based training compared to driver training on the road which is hardly used at this moment.

Are learners provided with gradually diminished coaching?

Currently/possibility: For the extrinsic feedback which is given by intelligent tutoring systems or human instructors the gradual decrease is implemented well, for example feedback is decreased after a certain number of successful task completions (DDS).

Principle number five, integration

Learning is promoted when learners are encouraged to integrate (transfer) new knowledge or skill into their everyday life.

Do learners have an opportunity to publicly demonstrate their new knowledge, reflect on their new knowledge, and create new ways to use their new knowledge?

Figure 1. Screenshot of the Drivemasters website (www.drivemasters.nl).

Currently/possibility: The simulator-based training should be integrated in a student's everyday life. A good example of how this can be done is making the driver training results publicly accessible. This is done by the Drivemasters portal, see Figure 1 (Drivemasters, 2007), where obtained training results are publicly visible and a competition element is added based on training results. Some training programs (ANWB, VR) focus more on the reflections of students after the lessons, using replays or discussion sessions.

4.4 Discussion

4.4.1 Summary of results and implications

An instructional design evaluation using Merrill's "first principles of instruction" revealed that all principles can be implemented on simulators and that there are many possibilities to improve the existing driver training programs. Although the instructional designs of the simulators that were investigated contain elements of the first principles of instruction, none of the simulators include all of the elements. The simulators could benefit from a didactical framework which is structurally used during the development of their individual lessons. Especially in the application phase, simulators are currently not only used to simulate the car and the environment, but also to simulate the complete real world driver training, including the feedback and instructions. Simulators have a lot of potential for improvement of training effectiveness by using intrinsic and extrinsic feedback mechanisms specifically designed for a certain training task. Because much more information concerning the state of the vehicle and the environment is available in a simulated environment compared to a real world environment, objective and valid performance measures could be used to present the student with feedback-on-performance which is not possible during real world driver training.

4.4.2 Reflections on the first principles of instruction

Merrill (2002a) extracted the first principles of instruction from some of the major instructional design models (e.g., Star Legacy, 4-MAT, Collaborative problem solving, Constructivist learning environments, 4C/ID Model, Learning by doing), of which many are included in the book of Reigeluth (1999) concerning instructional design theory. Merrill states: “As an instructional program implements more of the first principles of instruction, then there will be a corresponding increase in the quality and amount of learning that will occur.” This hypothesis is not empirically tested. Merrill assumes that if a principle is included in several instructional design theories, the principle has been found either through experience or empirical research to be valid.

The first principles of instruction show some resemblance to the well-known and widely adopted “nine events of instruction” as defined by Gagné (1985; 2005). These elaborate more on the presentation of feedback during application of newly acquired knowledge or skills. Concerning feedback, Merrill states in the diminishing-coaching corollary of principle number four that learning is promoted when learner are guided by appropriate feedback and coaching, including error detection and correction, and that learning is promoted when coaching is gradually withdrawn. In literature about driver training it is also noted that feedback-on-performance is essential for the learning of complex skills as driving (Groeger & Clegg, 2004). Skills are learned better and faster, if learners are given clear and immediate information of the effects of their actions on the measures which are used to analyze their performance (Romiszovski, 1999).

Both Gagné (2005) and Merrill (1997) stress the importance of an adaptation of the instructional design based on the objective learning outcome. The adaptations needed for the learning outcomes associated with car driving (i.e., procedure strategy, motor skills) are possible on simulators and are more difficult in real-world driving lessons. This is an advantage of simulator-based training that must be exploited further in order to strengthen the position of simulators in the driver training curriculum.

4.4.3 Conclusions and recommendations

The first principles of instruction provide a means to analyze the didactical quality of simulator-based driver training. Following the evaluation two conclusions can be drawn:

1. It is possible to implement all five first principles of instruction in simulator-based driver training programs.
2. None of the analyzed simulator-based training programs integrates all five principles in its program.

It is hypothesized that the didactical quality of training will improve if the principles of instruction are implemented in a systematic fashion in driver training simulators. It is recommended to use empirical experimentations to validate the use of the first principles of instruction to facilitate learning.

Chapter 5. The Effect of Concurrent Bandwidth Feedback on Learning the Lane-Keeping Task in a Driving Simulator

Abstract

The aim of this study was to investigate whether concurrent bandwidth feedback improves learning of the lane-keeping task in a driving simulator.

Previous research suggests that bandwidth feedback improves learning and that off-target feedback is superior to on-target feedback. This study aimed to extend these findings for the lane-keeping task.

Participants without a driver’s license drove five 8-min lane-keeping sessions in a driver training simulator: three practice sessions, an immediate retention session, and a delayed retention session 1 day later. There were four experimental groups (*n* = 15 per group): (a) on-target, receiving seat vibrations when the center of the car was within 0.5 m of the lane center; (b) off-target, receiving seat vibrations when the center of the car was more than 0.5 m away from the lane center; (c) control, receiving no vibrations; and (d) realistic, receiving seat vibrations depending on engine speed. During retention, all groups were provided with the realistic vibrations.

During practice, on-target and off-target groups had better lane-keeping performance than the nonaugmented groups, but this difference diminished in the retention phase. Furthermore, during late practice and retention, the off-target group outperformed the on-target group. The off-target group had a higher rate of steering reversal and higher steering entropy than the nonaugmented groups, whereas no clear group differences were found regarding mean speed, mental workload, or self-reported measures.

Off-target feedback is superior to on-target feedback for learning the lane-keeping task.

This research provides knowledge to researchers and designers of training systems about the value of feedback in simulator-based training of vehicular control.

De Groot, S., De Winter, J. C. F., López-García, J. M., Mulder, M., & Wieringa, P. A. (2011). The effect of concurrent bandwidth feedback on learning the lane keeping task in a driving simulator. *Human Factors, 53*, 50–62.

5.1 Introduction

Augmented feedback—that is, feedback other than the naturally available task-intrinsic feedback—is an important component of driver training. For example, learners may receive feedback on performance from a driving instructor (Hatakka et al., 2003) or automatically from a computer in case of simulator-based training (De Groot, De Winter, Mulder, & Wieringa, 2007). Nowadays, an estimated 150 driving simulators are used by driving schools in the Netherlands for basic driver training (SWOV Institute for Road Safety Research, 2010). Augmented feedback is a new trend in simulator-based driver training of tracking tasks, such as lane keeping and car following (Bekiaris, 2007).

Augmented feedback usually facilitates performance when provided during practice. However, it can have a negative impact on learning performance as measured during posttraining retention sessions without augmented feedback (Salmoni, Schmidt, & Walter, 1984; Schmidt & Wulf, 1997). For effective learning, augmented feedback should be provided sparingly and in such a way that the learner does not become dependent on it (see Swinnen, 1996, for a review). One possible way to accomplish this is to provide bandwidth feedback, which is feedback that depends on whether performance is within or outside a preset tolerance limit.

Bandwidth feedback has a long history starting with Thorndike's (1927) experiments in which participants received verbal right-wrong feedback after estimating the lengths of paper strips or drawing lines of particular lengths. Bandwidth feedback has also been used as concurrent feedback during tracking tasks. For example, Reynolds and Adams (1953) found that providing participants with an auditory click whenever on-target on a rotor pursuit task improved performance and learning (for reviews on motor learning, see Adams, 1964; Bilodeau & Bilodeau, 1961). With few exceptions, these early studies provided the augmented feedback when performance was on-target. Feedback was seen as a means of reinforcing habits and for providing motivation to learn. More recent studies used off-target feedback, providing the augmenting cues when performance deviated outside the tolerance limit (Goodwin & Meeuwsen, 1995; Lai & Shea, 1999; Lee & Carnahan, 1990; Lee & Maraj, 1994; Lintern, 1980; Reeve, Dornier, & Weeks, 1990; Sherwood, 1988; Smith, Taylor, & Withers, 1997).

One can argue whether on- or off-target feedback is most effective (Swinnen, 1996). On-target feedback may be more rewarding and stimulating for the learner. Off-target feedback, on the other hand, is less likely to distract the learner or make the learner dependent on the feedback, as no supplementary cues are provided when performing within reasonable standards (Lintern, 1991; Swinnen, 1996).

Only three studies could be found which made a direct comparison between on- and off-target feedback. Williams and Briggs (1962) let participants track a sinusoidal signal via a one-dimensional compensatory display. Augmented feedback was provided by means of auditory clicks when the participant was within or outside tolerance limits. During practice and immediate retention sessions without augmented feedback, the off-target group outperformed the on-target group, which in turn performed better than a control group not receiving augmented feedback during practice. A later study by Gordon and Gottlieb (1967) used a rotor pursuit task with visual augmented feedback from a light bulb. The on- and off-target groups outperformed the control group during practice and retention, with a slight and insignificant superiority for the off-target group. A more recent study by Cauraugh, Chen, and Radlo (1993) compared the effects of on- and off-target knowledge-of-results feedback on learning a 500-ms timing task. The reported effects were again in favor of off-target feedback, albeit not statistically significant. In summary, off-target feedback seems to be more effective than on-target feedback, although the reported effects are not uniformly supportive.

This study focuses on lane keeping, which is a fundamental control task that drivers should master to drive safely. Crashes in traffic are not usually the result of adverse perceptual motor skills of lane keeping per se but are generally the consequence of a complex interplay of events, including, for example, inattention lapses and intentional violations. Nonetheless, lane-keeping performance has been associated with driver impairment and has been used as a proxy variable for road safety in numerous studies (e.g., Brookhuis & De Waard, 1993; Ranney, Harbluk, & Noy, 2005; Verster, Veldhuijzen, Patat, Olivier, & Volkerts, 2006).

Our aim was to compare the effects of on-target versus off-target feedback for learning the lane keeping task in a driving simulator. We investigated concurrent feedback: a nondirectional seat vibration indicating in real time whether the center of the car was inside or outside an invisible 1-m-wide band during practice sessions. The lane-keeping task features clear intrinsic feedback, which reduces the risk of learners becoming dependent on augmenting cues (cf. Kinkade, 1963). We chose vibrotactile augmented feedback to minimize interference with the visual driving task. Vibrotactile feedback has been used for effective presentation of in-vehicle warning and navigation signals (Ho, Reed, & Spence, 2007; Jones & Sarter, 2008; Lee, Hoffman, & Hayes, 2004; Suzuki & Jansson, 2003; Van Erp & Van Veen, 2004).

After three practice sessions, learning performance was assessed during an immediate retention session and a 1-day delayed retention session, considering the role of sleep in performance consolidation (Walker, Brakefield, Morgan, Hobson, & Stickgold, 2002). Earlier research has indicated that information-processing abilities and mental effort have an important influence on learning (Ackerman, 1987; Guadagnoli & Lee, 2004; Lee, Swinnen, & Serrien, 1994). Therefore, we measured these constructs using a secondary task and a questionnaire.

5.2 Method

5.2.1 Participants

We recruited 60 persons without a driver's license from the Delft University of Technology student community. The participants completed an intake questionnaire with the following items: (a) gender (male or female), (b) age, (c) "Did you take driving lesson(s) on the road?" (yes or no), (d) "Have you ever driven in a simulator before?" (yes or no), (e) "Do you play video games for at least 1 hour per week?" (yes or no), and (f) "I have good steering skills (for example in cycling or computer games)" (on a 10-point scale from 1 = completely disagree to 10 = completely agree). Earlier research suggested that these items are predictive of lane-keeping performance in a simulator (Backlund, Engström, Johannesson, & Lebram, 2010; De Winter et al., 2006, 2009). Of the 60 participants, 13 were female, 23 had started taking driving lessons on the road, 6 had driven in a simulator before, and 31 reported playing video games for more than 1 hr per week. The mean age was 19.7 years (SD = 2.3), and the mean self-reported steering skill was 6.8 (SD = 1.6).

5.2.2 Apparatus

The driving simulator (Figure 1) was fixed base and provided a realistic simulation of a middle-class passenger car with 180° field of view and surround sound. This simulator is used for initial driver training in the Netherlands (Green Dino, 2008). The pedals, steering wheel, ignition key, and seat resembled those of an actual car, and gear changing was automatic. The steering wheel provided force feedback using a DC torque motor coupled to the steering shaft through an 8:1 gearbox (Dunkermotoren, Bonndorf, Germany). The applied torque was a function of the speed and steering wheel angle from the car model, representing static friction and aligning torque.

The virtual world was projected by means of three LCD projectors (front projector NEC VT676, brightness 2,100 ANSI lumens, contrast ratio 400:1, resolution 1,024 × 768 pixels; side projectors NEC VT470, brightness 2,000 ANSI lumens, contrast ratio 400:1, resolution 800 × 600 pixels), and the dashboard, interior, and mirrors were integrated in the projected image. Vibrations were provided by a tactile transducer (ShakerCentre, Brentwood, UK) fixed vertically onto the metal frame of the seat.

5.2.3 Experimental groups

There were four experimental groups: on-target, off-target, control, and realistic (n = 15 per group). Assignment of a new participant to one of the four groups was determined so as to minimize the differences between the groups in terms of number of women and mean age.

The on-target and off-target groups practiced with concurrent vibratory bandwidth feedback. For the on-target group, the vibration automatically switched on when the center of the car was within 0.5 m of the center of the lane. For the off-target group, the vibration was operative when the center of the car was more than 0.5 m away from the lane center.

Figure 1. Driving simulator during the experiment. The dashboard configuration shows the speed and engine rpm. Gear changing was automatic.

The vibrations applied for the on-target and off-target groups had a frequency of 100 Hz and a constant intensity of 0.09 m/s^2 root mean squared (RMS), values chosen subjectively by the experimenters so that the vibrations were clearly perceivable but not annoying. The vibration intensity was measured during pilot sessions with a vertical accelerometer sampled at 200 Hz, positioned on the seat cushion with a human sitting on top. Note that the augmented feedback was nondirectional; it indicated whether the participant was on-target or off-target, not whether left or right steering was required. The control group received no vibration feedback during practice. Participants in the realistic condition drove with vibrations depending on the speed of the simulated engine. The realistic vibrations represented a combination of the first- and second-order vibration modes of a four-cylinder engine. The vibration intensity varied between 0.06 m/s^2 and 0.12 m/s^2 RMS.

As a comparison, we measured RMS acceleration in a car with diesel engine while standing still and varying engine speed between 1,000 rpm and 3,500 rpm. The obtained RMS values were in approximately the same range (0.04 m/s^2 to 0.12 m/s^2). Paddan and Griffin (2002) measured RMS on the seats of 25 cars traveling at different speeds on tarmac and concrete. The RMS ranged between 0.16 m/s^2 and 0.78 m/s^2. Pielemeier, Jeyabalan, Meier, and Otto (1997) measured just noticeable differences for three trained persons exposed to vibrations on a car seat. For a reference vibration of 0.08 m/s^2 RMS, perceptual thresholds ranged between 0.006 m/s^2 and 0.018 m/s^2. The realistic vibrations were therefore clearly perceivable and corresponded to those in a real car.

5.2.4 Procedure and tasks

After recruitment, an e-mail was sent to the participants containing the time and location of the experiment as well as a protocol explaining that the task goal was to drive perfectly in the center of the right lane. All participants signed an informed consent form.

Participants first performed an 8-min baseline session in the simulator to practice the auditory reaction time task (described in this section). Next, five 8-min driving sessions were completed: three practice sessions, an immediate retention session, and a delayed retention session on the following day. In the retention phase, all participants were provided with the realistic vibrations. In other words, on-target, off-target, and control groups transferred to a higher-fidelity configuration, whereas the realistic group was provided with the same vibrations during practice and retention.

All driving sessions took place on a two-lane 7.5-km lap in a country environment without intersections or other vehicles. The lap contained 11 right curves and 10 left curves of 90 degrees, for which braking was needed. The lane width was 5 m. The road contained a tunnel and two 4-m hills. The road surface was uniform and flat; there was no horizontal curvature of the road profile as can be found in other simulators (Allen et al., 1999). Like in a real car, small steering corrections were continuously needed on the straights to keep the car in the center of the lane. Previous research on this simulator showed that performance on lane-keeping and steering tasks predicted the chance of passing the Dutch driving license test (De Winter et al., 2009).

Before each driving session, a series of textual instructions were projected in the simulator, explaining how to obtain a proper seating position, how to perform the reaction time task, how to control the car (steering wheel, throttle, and brake pedals), and, finally, the goal of the task (to drive in the center of the lane) and (only when appropriate) how to interpret the augmented feedback.

During both the baseline session and the five driving sessions, an auditory reaction time task had to be performed. Earlier driving simulator research demonstrated that reaction time tasks were sensitive in detecting differences in workload as a function of tactile feedback during car following (Mulder et al., 2004) and complexity of the driving context (Cantin, Lavallière, Simoneau, & Teasdale, 2009). The task was to react as quickly as possible to a 0.1-s beeping sound produced at a random time interval between 4 s and 8 s. The reaction time was measured from the moment the beep was produced until the moment the participant pressed the horn (central piece on the steering wheel). After pressing the horn, a second beep, with lower tone, was produced as a confirmation. No confirmation sound was produced when the participant did not react within 2 s. After each of the six sessions, participants were asked to step out of the simulator to complete the Rating Scale Mental Effort (Zijlstra, 1993).

5.2.5 Dependent measures

The following dependent measures were calculated for each practice and retention session.

Percentage on-target. This variable was the percentage of driving time that the center of the vehicle was within 0.5 m of the center of the right lane of the two-lane road. Percentage on-target was our primary performance measure, as it reflects the behavior that was trained with the bandwidth feedback. It was also used in the seminal articles on off-target versus on-target feedback (Gordon & Gottlieb, 1967; Williams & Briggs, 1962). Data from 10 s prior to 20 s after a road departure were excluded for calculating percentage on-target, mean speed, steering reversal rate, steering entropy, and root mean squared error (RMSE) lane center.

Mean speed. This variable was mean speed of the simulated vehicle. Mean speed is an indicator of the participant's efficiency of proceeding along the course.

Number of departures. This variable measured the number of occasions that the vehicle was outside the lane boundaries with all its edges. The distance of the center of the vehicle to the center of the right lane was dependent on the heading of the vehicle and was on average 3.93 m (SD = 0.36 m) and 9.27 m (SD = 0.46 m) for departures to the right and to the left, respectively. Road departures were typically the consequence of improper lane-keeping behavior or loss of control because of approaching a curve too fast. When a road departure occurred, the participant was automatically placed back on the center of the right lane with zero speed.

Steering reversal rate. The number of steering wheel reversals per second is a measure of control activity that does not necessarily correlate with absolute measures of lane-keeping performance (McLean & Hoffmann, 1975). A high steering reversal rate indicates that the driver made many steering corrections and gave high-frequent steering wheel input. A reversal was defined as a change from a clockwise movement to a counterclockwise movement, provided that the counterclockwise steering velocity exceeded 3.0 deg/s (Theeuwes, Alferdinck, & Perel, 2002). Previous research has shown that the steering reversal rate is sensitive to changes in task demand and control effort, albeit in a complex manner (MacDonald & Hoffmann, 1980).

Steering entropy. Steering entropy, representing the information content and smoothness of the steering wheel angle, was calculated as described by Nakayama, Futami, Nakamura, and Boer (1999), except that a 4-Hz instead of a 7-Hz resample frequency was applied. This modification was recommended by Boer, Rakauskas, Ward, and Goodrich (2005) to make the entropy measure more sensitive to changes of the driving task. The bin width was set at 2.15°, representing 60% of the frequency distribution of the prediction error averaged across all sessions and participants (60% was also taken from Boer et al., 2005). High steering entropy means that the driver's steering wheel input was discontinuous and irregular.

RMSE lane center. This variable refers to the RMSE of the distance of the center of the vehicle to the center of the right lane. It describes how accurately the driver kept the vehicle near the lane center but in a way that is more remotely related to the bandwidth feedback than percentage on-target.

Mean reaction time. The mean reaction time on the auditory reaction time task, representing the participant's mental workload, was also calculated. Reaction times less than 0.15 s (anticipatory responses or "false alarms") and reaction times exceeding 2 s (missed responses) were excluded.

Effort. The Rating Scale Mental Effort provides an indication of the participant's level of effort expenditure (Zijlstra, 1993). This rating scale was presented on an A4 paper with a 150-mm vertical bar with anchors every nine points from 2 mm (*absolutely no effort*) to 112 mm (*extreme effort*). The instructions on the form stated, "Please indicate, by marking the vertical axis below, how much effort it took for you to complete the task you have just finished."

After finishing the delayed retention task, participants completed a 14-item questionnaire. The items represented statements to be answered on a scale from 1 (*completely disagree*) to 10 (*completely agree*) concerning the participant's enjoyment, task difficulty, simulator sickness, concentration, realism of the simulator, the ability to judge speed and distance, and the need for feedback about task performance. In this study, we were interested in the response to the statement "Keeping the car in the center of the lane was easy" as an indicator of workload. The obtained scores were reversed so that a score of 1 corresponded to *easy* and a score of 10 corresponded to *difficult*.

5.2.6 Statistical analyses

Comparisons between groups were conducted using a full-factorial analysis of covariance (ANCOVA). ANCOVA was performed for the four groups together and for the following four group pairs: on-target versus the nonaugmented groups (control and realistic combined), off-target versus the nonaugmented groups, on-target versus off-target, and control versus realistic. *F* statistics comparing group pairs were converted to Cohen's *d* (i.e., the standardized mean difference) for ease of interpretation, as d conveys information about the sign and size of the effect independent of the sample size.

Comparisons between sessions were carried out with repeated-measures ANCOVA. To assess changes during practice, Practice 1 and Practice 3 were included as within-subjects variables. To assess changes from practice to retention, Practice 3 and delayed retention were included as within-subjects variables.

The covariate used in ANCOVA was the participants' initial aptitude. This variable was calculated from the following information acquired prior to commencing the first practice session: the six items from the intake questionnaire and the mean reaction time and effort of the baseline session. The matrix of these variables (60 participants × 8 variables) was reduced into one score, representing the initial aptitude, by taking the first principal component based on the correlation matrix. The mean initial aptitude scores were .05, −.28, .19, and .04 for on-target, off-target, control, and realistic groups, respectively. These means were not significantly different as determined with a one-way ANOVA (F = 0.57, p = .640).

5.3 Results

An investigation of the normal probability plots showed that the number of departures had a considerably skewed distribution. Therefore, a square root transformation was applied on this measure. The correlation matrix in Table 1 shows that the percentage on-target, the number of departures, and RMSE lane center were substantially correlated, as these are all measures of lane-keeping performance. Furthermore, steering reversal rate and steering entropy turned out to be strongly correlated (.92), which implies that these variables essentially describe the same construct.

A low reversal rate and low entropy are indicative of “smoothness of control” (e.g., Nakayama et al., 1999, p. 2). The initial aptitude covariate was significantly predictive of five of the nine dependent measures. The reliabilities, shown on the diagonal of Table 1, were mostly above .8 and were considered adequate for ergonomics research (cf. Liu & Salvendy, 2009). The number of departures had a relatively low reliability (.56), which implies that significant effects are less likely for this measure. The low reliability can be explained by the fact that departures were relatively infrequent events, with 32%, 67%, 68%,80%, and 75% of the participants having no departures during the five respective driving sessions.

Table 2 shows the uncorrected means and standard deviations of the dependent measures per group and session, Table 3 shows the group effects and the within-subjects effects of the four groups together, and Table 4 shows the results of the comparisons between group pairs.

Table 1. Correlation matrix amongst the dependent measures (N = 60) with test-retest reliabilities on the diagonal.

	1	2	3	4	5	6	7	8	9
1. Percentage on target	.87								
2. Mean speed (m/s)	−.07	.88							
3. Number of departures	−.56	.29	.56						
4. Steering reversal rate (#/s)	.00	.49	.23	.85					
5. Steering entropy	−.08	.57	.23	.92	.88				
6. RMSE lane center (m)	−.85	−.01	.63	.08	.08	.82			
7. Mean reaction time (s)	−.31	−.17	.24	.00	.12	.31	.92		
8. Effort	.00	−.32	.11	−.15	−.14	.15	.17	.91	
9. Task difficulty	−.31	−.25	.15	−.07	−.02	.31	.45	.30	
10. Initial aptitude	.42	.35	−.25	.07	−.02	−.40	−.49	.05	−.49

The correlations were determined by first averaging the measures across the five driving sessions. Group effects were eliminated by subtracting the group mean. Correlations of magnitude greater than or equal to .26 are significant, $p < .05$. The test-retest reliabilities were calculated by taking the correlation (controlling for group effects) between Practice 2 and Practice 3.

Table 2. Uncorrected means per dependent measure, experimental group, and driving session (standard deviations between parentheses).

	Practice			Retention	
	1	2	3	Immediate	Delayed
Percentage on-target					
On-target	67 (10)	71 (11)	73 (10)	70 (11)	67 (10)
Off-target	66 (11)	76 (9)	80 (8)	76 (12)	74 (16)
Control	56 (13)	62 (14)	63 (14)	66 (14)	69 (10)
Realistic	56 (9)	64 (9)	64 (11)	66 (15)	69 (8)
Mean speed (m/s)					
On-target	16.3 (2.3)	15.6 (2.2)	15.4 (1.8)	15.6 (2.0)	16.4 (2.2)
Off-target	16.4 (2.3)	15.8 (1.9)	16.0 (2.0)	15.6 (1.9)	15.9 (2.3)
Control	16.1 (1.1)	16.0 (1.7)	16.6 (1.7)	16.6 (1.7)	16.7 (1.3)
Realistic	16.2 (1.6)	16.0 (1.4)	16.4 (1.6)	15.9 (1.1)	16.4 (1.5)
Number of departures					
On-target	1.87 (1.85)	0.47 (1.06)	0.33 (0.62)	0.07 (0.26)	0.27 (0.46)
Off-target	1.33 (1.23)	0.60 (0.63)	0.53 (0.99)	0.33 (0.62)	0.53 (0.99)
Control	1.67 (1.63)	0.40 (0.83)	0.40 (0.51)	0.60 (0.91)	0.13 (0.35)
Realistic	1.53 (1.85)	0.80 (1.42)	0.53 (0.92)	0.20 (0.56)	0.53 (0.92)
Steering reversal rate (#/s)					
On-target	0.58 (0.087)	0.53 (0.080)	0.54 (0.091)	0.53 (0.094)	0.54 (0.071)
Off-target	0.63 (0.069)	0.58 (0.061)	0.58 (0.070)	0.56 (0.056)	0.55 (0.055)
Control	0.56 (0.094)	0.51 (0.083)	0.50 (0.076)	0.49 (0.082)	0.50 (0.067)
Realistic	0.58 (0.050)	0.54 (0.068)	0.52 (0.054)	0.51 (0.061)	0.52 (0.036)
Steering entropy					
On-target	0.93 (0.049)	0.90 (0.043)	0.90 (0.047)	0.90 (0.047)	0.90 (0.045)
Off-target	0.95 (0.025)	0.92 (0.036)	0.92 (0.030)	0.90 (0.030)	0.90 (0.029)
Control	0.90 (0.039)	0.89 (0.044)	0.88 (0.046)	0.88 (0.044)	0.89 (0.035)
Realistic	0.92 (0.032)	0.90 (0.045)	0.89 (0.043)	0.89 (0.035)	0.89 (0.027)
RMSE lane center (m)					
On-target	0.82 (0.29)	0.67 (0.27)	0.63 (0.27)	0.63 (0.23)	0.63 (0.24)
Off-target	0.81 (0.26)	0.57 (0.15)	0.52 (0.17)	0.53 (0.22)	0.56 (0.22)
Control	0.90 (0.33)	0.73 (0.27)	0.64 (0.20)	0.63 (0.21)	0.57 (0.12)
Realistic	1.01 (0.26)	0.74 (0.17)	0.68 (0.20)	0.60 (0.18)	0.64 (0.16)
Mean reaction time (s)					
On-target	0.78 (0.12)	0.69 (0.11)	0.70 (0.12)	0.67 (0.12)	0.64 (0.12)
Off-target	0.84 (0.17)	0.76 (0.14)	0.74 (0.14)	0.72 (0.14)	0.67 (0.14)
Control	0.79 (0.14)	0.69 (0.13)	0.66 (0.12)	0.64 (0.11)	0.63 (0.11)
Realistic	0.79 (0.15)	0.71 (0.13)	0.69 (0.14)	0.65 (0.11)	0.64 (0.11)
Effort					
On-target	71 (24)	66 (22)	63 (19)	67 (20)	61 (18)
Off-target	69 (24)	63 (23)	61 (26)	58 (25)	58 (25)
Control	81 (20)	75 (20)	74 (21)	71 (18)	66 (19)
Realistic	67 (18)	63 (24)	57 (25)	57 (28)	50 (17)

Table 3. Group effects and within-subjects effects (*F*-values; *p*-values in parentheses).

	Practice *F*(3,55)	Immediate Retention *F*(3,55)	Delayed Retention *F*(3,55)	Practice 1 vs. Practice 3 *F*(1,55)	Practice 3 vs. Delayed Retention *F*(1,55)
Percentage on-target	8.70 (.000)	4.18 (.010)	1.87 (.145)	47.1 (.000)	0.09 (.771)*
Mean speed	0.29 (.830)	1.06 (.374)	0.18 (.911)	0.57 (.454)	1.46 (.233)
Number of departures	0.05 (.984)	1.72 (.173)	0.70 (.555)	37.0 (.000)	0.75 (.389)
Steering reversal rate	2.63 (.059)	2.22 (.096)	2.63 (.059)	60.4 (.000)	1.34 (.251)
Steering entropy	2.50 (.069)	1.09 (.361)	0.88 (.457)	40.9 (.000)	1.92 (.172)
RMSE lane center	2.52 (.068)	1.91 (.138)	0.97 (.412)	76.0 (.000)	0.64 (.429)
Mean reaction time	0.38 (.766)	0.82 (.490)	0.11 (.956)	85.9 (.000)	25.0 (.000)
Effort	1.39 (.257)	1.12 (.348)	1.64 (.190)	15.4 (.000)	3.56 (.065)

*There was a significant Session x Group interaction, $F(3,55) = 5.35$, $p = .003$.

Table 4. Comparisons between group pairs (Cohen's *d*).

	Practice			Retention	
	1	2	3	Immediate	Delayed
On-target (*n* = 15) vs. nonaugmented groups (*n* = 30)					
Percentage on target	1.00**	0.81*	0.86**	0.41	−0.15
Mean speed	0.10	−0.21	−0.60#	−0.38	−0.08
Number of departures	0.08	−0.17	−0.21	−0.49	−0.02
Steering reversal rate	0.10	−0.02	0.38	0.38	0.46
Steering entropy	0.30	0.16	0.29	0.27	0.29
RMSE lane center	−0.45	−0.29	−0.20	0.05	0.12
Mean reaction time	−0.14	−0.07	0.18	0.16	0.06
Effort	−0.12	−0.12	−0.09	0.13	0.13
Off-target (*n* = 15) vs. nonaugmented groups (*n* = 30)					
Percentage on target	0.97**	1.23***	1.54***	1.01**	0.58#
Mean speed	0.36	0.05	−0.13	−0.30	−0.20
Number of departures	−0.14	0.14	−0.10	−0.03	0.15
Steering reversal rate	0.81*	0.66*	0.94**	0.81*	0.84**
Steering entropy	0.94**	0.63*	0.76*	0.55#	0.53#
RMSE lane center	−0.53#	−0.85**	−0.90**	−0.63*	−0.37
Mean reaction time	0.16	0.38	0.34	0.47	0.17
Effort	−0.24	−0.27	−0.20	−0.20	0.04
Off-target (*n* = 15) vs. On-target (*n* = 15)					
Percentage on target	0.11	0.65	0.98*	0.90*	0.88*
Mean speed	0.25	0.29	0.50	0.21	−0.06
Number of departures	−0.23	0.31	0.01	0.54	0.16
Steering reversal rate	0.77#	0.81*	0.51	0.41	0.32
Steering entropy	0.54	0.53	0.50	0.31	0.17
RMSE lane center	−0.13	−0.59	−0.73#	−0.90*	−0.50
Mean reaction time	0.31	0.48	0.17	0.30	0.08
Effort	−0.14	−0.21	−0.22	−0.44	−0.09
Realistic (*n* = 15) vs. Control (*n* = 15)					
Percentage on target	0.05	0.20	0.14	0.05	0.06
Mean speed	0.12	−0.03	−0.13	−0.55	−0.22
Number of departures	−0.26	0.26	0.01	−0.55	0.55
Steering reversal rate	0.32	0.49	0.23	0.28	0.42
Steering entropy	0.54	0.36	0.27	0.19	0.18
RMSE lane center	0.37	0.03	0.16	−0.22	0.44
Mean reaction time	−0.07	0.07	0.27	0.04	0.00
Effort	−0.74#	−0.58	−0.74#	−0.56	−0.93*

***$p < .001$, **$p < .01$, *$p < .05$, #$p < .1$.

Percentage on-target. During practice, on-target and off-target groups drove a significantly larger percentage of the time on the lane center target than the nonaugmented groups. For on-target drivers, this advantage disappeared in the retention phase. For off-target drivers, much of the advantage remained during immediate retention, although a clear drop in performance was evident during delayed retention. The differences between off-target and the nonaugmented groups were d = 1.01 (p = .002) and d = 0.58 (p = .069) during immediate and delayed retention, respectively, which classify as moderate to strong effects (Cohen, 1988).

As expected, drivers in the off-target condition outperformed those in the on-target condition. This difference was not apparent during Practice 1 but appeared in Practice 2, reached significance in Practice 3, and persisted during immediate and delayed retention. There was no significant difference between control and realistic groups. Percentage on-target increased for all groups during practice, a strong effect clearly illustrated in Figure 2. The groups driving without augmented feedback continued to improve from Practice 3 to delayed retention, $F(1, 27) = 11.5$, $p = .002$, whereas the performance of the augmented feedback groups deteriorated, $F(1, 27) = 7.10$, $p = .008$.

Mean speed. There were no significant group differences during practice or retention concerning the mean speed. Table 3 shows that there were no significant differences between Practice 1 and 3 or between Practice 3 and delayed retention. Nonetheless, there were significant differences between the sessions when considering the five sessions combined, $F(4, 220) = 3.29$, $p = .012$. A post hoc analysis indicated that the speed during Practice 2 was significantly slower than during Practice 1, Practice 3, and delayed retention. Furthermore, participants drove significantly faster during delayed retention as compared with immediate retention. As can be seen in Table 2, the effects were small on an absolute scale: All mean speeds of the 20 combinations of session and group ranged between 15.4 m/s and 16.7 m/s.

Number of departures. No significant group differences for the number of road departures were found. Control and off-target groups had a slightly elevated

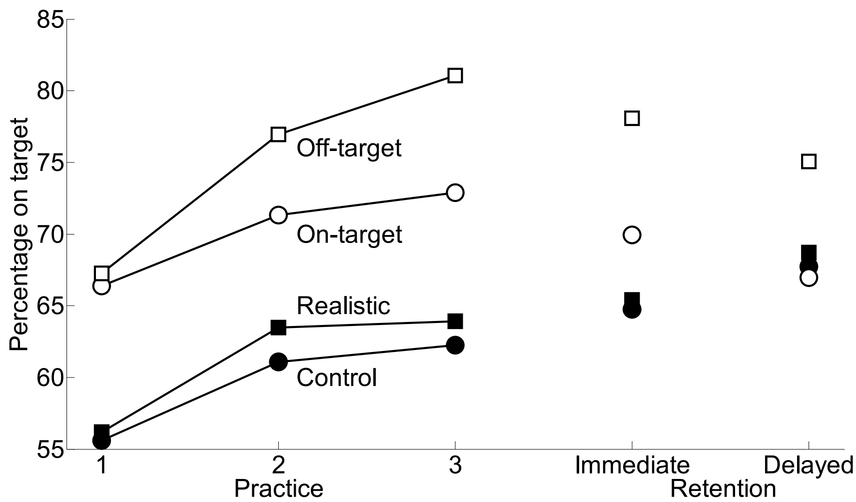

Figure 2. Estimated marginal means of the percentage on target for the four experimental groups during practice and retention. Estimated marginal means are the predicted means by holding the initial aptitude covariate at its mean value of zero.

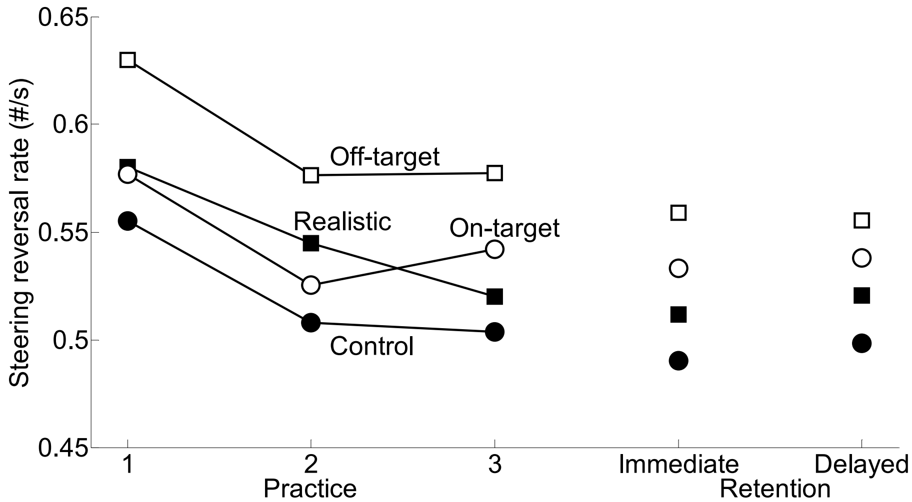

Figure 3. Steering reversal rate for the four experimental groups during practice and retention. Estimated marginal means are the predicted means by holding the initial aptitude covariate at its mean value of zero.

number of departures during immediate retention, a statistically insignificant effect. The number of departures decreased from Practice 1 to Practice 3 but did not decrease significantly from Practice 3 to delayed retention.

Steering reversal rate and steering entropy. The results in Table 3 show that participants' steering became significantly smoother from Practice 1 to Practice 3. Table 4 shows that drivers in the off-target condition had less smoothness of control (i.e., higher reversal rate and entropy) than the nonaugmented groups during practice and retention with all *p* values below .05, except for steering entropy during retention. The results for the steering reversal rate are illustrated in Figure 3.

RMSE lane center. This measure correlated strongly with the number of departures and percentage on-target. RMSE lane center strongly improved with practice. The same trends between groups were found as for percentage on-target, albeit with weaker effects.

Mean reaction time. There were no significant group differences, but there was a clear learning effect: The reaction time decreased from Practice 1 to Practice 3 and decreased further from Practice 3 to delayed retention.

Effort. No significant group differences were found. Drivers in the control group tended to report higher effort than those in the realistic condition. There was a clear learning effect: Participants provided lower effort ratings for Practice 3 as compared with Practice 1 and tended to further reduce effort from Practice 3 to delayed retention. ANCOVA performed on each of the 14 post-experiment questions separately revealed not one significant difference between the four groups. The marginal means of the task difficulty scores were 5.2, 5.2, 5.2, and 4.4 on a scale from 1 to 10 for on-target, off-target, control, and realistic groups, respectively (F = 0.910, p = .442).

5.4 Discussion

This study investigated the potential of concurrent vibratory bandwidth feedback for learning the lane-keeping task in a driving simulator. The on-target and off-target groups outperformed the control and realistic groups during practice and immediate retention, but consistent with the guidance hypothesis (Salmoni et al., 1984), performance of the augmented-feedback groups deteriorated during retention. The most significant finding of this research is that performance of drivers in the off-target condition was superior to that of drivers in the on-target condition during practice and retention. This replicates the early work of Williams and Briggs (1962) for an ecologically valid task. No clear group differences were found regarding mean speed, mental workload, or self-reported measures.

Literature describes three possible reasons that off-target feedback may be superior to on-target feedback. First, off-target feedback is sensitive to the needs of the learner because the amount of feedback is smaller for the better skilled drivers and reduces when performance improves (T. D. Lee et al., 1994). Williams and Briggs (1962) used similar argumentation to explain why the off-target group performed better than the on-target group during retention trials: “This is logical in view of the similarity between the training and transfer trials for the off-target group; during training these subjects experienced a diminution in number of auditory clicks as result of increasing skill” (p. 524). Indeed, during Practice 1, on-target and off-target groups on average received vibrations 66% and 33% of the time, respectively; during Practice 3, this difference had grown to 73% and 18%, respectively.

A second reason drivers receiving on-target feedback underperformed may be that because the augmented feedback is provided when they are within the tolerance limit, it blocks processing of intrinsic visual feedback (Lintern, 1980). Off-target feedback, on the other hand, does not distract from intrinsic feedback sources and does not permit the learner to become dependent on the augmented feedback.

A third advantage of off-target feedback may be that it stimulates the learner to correct performance only when errors are large, preventing maladaptive short-term corrections beyond the precision of the motor system (Lee & Carnahan, 1990; Sherwood, 1988).

However, these reasons cannot satisfactorily explain the superior retention performance of the off-target group, because on-target and off-target groups in essence used an identical on-off feedback algorithm, switching sign on bandwidth crossings; only the sign of the feedback was opposite. Because of the identical algorithm, there was no apparent reason why learners should become more dependent on on-target feedback any more than on off-target feedback. Furthermore, our comparison between realistic and control groups showed that vibrations per se did not have a negative effect on practice and retention performance (for discussion on the effect of disturbance motion cues, see Caro, 1979). Moreover, both on-target and off-target groups transferred to a new condition (the realistic vibrations), so drivers receiving off-target feedback were not more familiar with the retention condition than those receiving on-target feedback.

We conducted a supplementary analysis of steering correction behavior. In the practice phase, and particularly during Practice 3, drivers in the off-target group drove on average shorter periods outside the bandwidth than those in the on-target group (M = 1.78 s vs. M = 2.28 s, p = .011, d = −1.04), who in turn drove shorter periods outside the band than drivers in the nonaugmented groups (M = 2.95 s). In other words, drivers in the off-target group made more rapid steering corrections than those in the on-target group, which is in line with a commonly observed finding in psychophysics called the onset advantage. As explained by Fisher and Miller (2008), research suggests that the sudden onset of a stimulus is a more powerful perceptual event than a stimulus offset, facilitating low-level perceptual processing and resulting in faster reaction times.

We also investigated the number of inside-to-outside bandwidth crossings. The on-target group performed equivalently to the nonaugmented groups throughout the experiment. The off-target group, on the other hand, started worst and gradually improved throughout practice and retention. Eventually, during delayed retention, the off-target group had considerably fewer bandwidth crossings than did the on-target group (M = 40.6 vs. M = 55.1, p < .001, d = −1.43). This indicates that drivers in the off-target group had learned to avoid errors that were previously associated with vibration onsets.

Concluding, the difference in performance between on-target and off-target groups seems to be attributable to fundamental differences in the way humans process information when a signal switches from on to off or vice versa. Together with the earlier work on the same topic (Cauraugh et al., 1993; Gordon & Gottlieb, 1967; Williams & Briggs, 1962), it is now established that for effective learning of tracking skills, the onset of a stimulus should be associated with erroneous performance, not with correct performance. Our results also suggest that lane departure warning systems that become increasingly available in real vehicles may have a benefit for learning.

The off-target group had a higher steering reversal rate and steering entropy than did the nonaugmented groups. This lower level of steering smoothness was present during practice and extended to the retention sessions, indicating that it was something that was learned and that persisted. Less steering smoothness for the off-target group is in agreement with the finding discussed already, that the participants from the off-target group reacted more strongly to lane center errors. Indeed, a moderately lower level of steering smoothness could mean that the driver is controlling the car more actively. On the other hand, less smoothness may be indicative of inadequate anticipation of the curves ahead.

What constitutes optimal steering smoothness is clearly an interesting topic of further research. Using bandwidth feedback based on a predicted lane center error instead of the momentary lane center error may be a means to stimulate the learning of anticipatory steering behavior.

In this study, participants were trained to drive close to the center of the lane. It is acknowledged that driving around curves can be successfully accomplished in other ways than centering the vehicle in the lane. For example, approaching the inside of the lane may be considered a preferred driving style. We are planning to investigate the effectiveness of off-target feedback in our racing simulator with the bandwidth feedback centered on the ideal racing line. Furthermore, we hypothesize that the effectiveness of concurrent off-target bandwidth feedback is generalizable and can also be applied to the training of longitudinal control to keep the speed of the car within specified limits. The present experiment featured 24 min of practice per participant and thus provided an indication of initial learning only. Asymptotic driving skill is normally obtained after months or even years of experience (Mayhew, Simpson, & Pak, 2003). It is therefore recommended to investigate the longer-term effects of augmented feedback. Another recommendation is to investigate the potential of augmented feedback using more complex driving tasks, such as negotiating intersections.

Chapter 6. The effect of tire grip on learning driving skill and driving style: A driving simulator study

Abstract

There is a need for training methods that improve the driving skill and driving style of novice drivers. Previous research in motor learning has shown that degrading the task conditions during practice can enhance long-term retention performance. Inspired by these findings, this study investigated the effects of the tire-road friction coefficient on learning a self-paced lane-keeping task in a driving simulator. A sample of 63 young and inexperienced drivers were divided into three groups, low grip (LG), normal grip (NG), and high grip (HG), who practiced driving with a friction coefficient of 0.45, 0.90, and 1.80, respectively. All groups drove six 8 min sessions on a road with curves in a rural environment: four practice sessions, an immediate retention session, and a delayed retention session on the next day. The two retention sessions were driven with normal-grip tires. The results show that LG drove with lower speed than NG during practice and retention. Transferring from the last practice session to the immediate retention session, LG's workload decreased, as measured with a secondary task, whereas HG's workload increased. During the immediate retention session, LG had less road departures than HG, but HG drove closer to the lane center in curves than the other two groups. HG reported elevated confidence during practice, but not in retention. In conclusion, this simulator-based study showed that practicing with low-grip tires resulted in lower driving speeds during retention tests, an effect which persisted overnight. These results have potential implications for the way drivers are trained.

De Groot, S., Centeno Ricote, F., & De Winter, J. C. F. (2012). The effect of tire grip on learning driving skill and driving style: A driving simulator study. *Transportation Research Part F, 15*, 413–426.

6.1 Introduction

It is widely established that young novice drivers are overrepresented in motor vehicle crashes, a major public health concern (McCartt, Mayhew, Braitman, Ferguson, & Simpson, 2009; Williams, 2006). The young driver problem is complex, but in simple terms, driving skill and driving style can be seen as the main contributing factors (Elander, West, & French, 1993). The distinction between driving skill and driving style resembles the distinction between lower-order and higher-order driving skills, and errors and violations (Harrison, 1999; Hatakka, Keskinen, Gregersen, Glad, & Hernetkoski, 2002; Reason, Manstead, Stradling, Baxter, & Campbell, 1990). Driving skill reflects the way in which a person is able to drive. Newly licensed drivers tend to have an elevated mental workload and inefficient visual search, hazard perception, and vehicle control abilities (Crundall, Underwood, & Chapman, 1999; Drummond, 1989; Falkmer & Gregersen, 2005; Lee, 2007; McKnight & McKnight, 2003; Pradhan et al., 2005; Vlakveld, 2011). Driving style is the way in which a driver chooses to drive and is governed by a combination of social, neurobehavioral, and biological mechanisms (Dahl, 2008; Evans, 2006). Young drivers, males in particular, have a high willingness to take risks and tend to be overconfident in their own abilities (Clarke, Ward, & Truman, 2005; Finn & Bragg, 1986; Gregersen, 1996; Jonah, 1986; Simons-Morton, Lerner, & Singer, 2005).

Classic training methods that mainly target driving skills, such as skid-control training or basic driver training, appear to be ineffective and may even inflate crash risk (Elvik & Vaa, 2004; Katila, Keskinen, & Hatakka, 1996; Lund, Williams, & Zador, 1986). Recommendations have been made to target driving style during young driver training, for example, through group discussions and self-reflection (Hatakka et al., 2002). Even though driver training and licensing methods are being continuously revised (Twisk & Stacey, 2007), the young driver problem has remained and there is a need to explore new methods of training to improve the driving skills as well as the driving style of young novice drivers. Interactive driving simulators and PC-based programs are recognized as potentially effective tools in driver training (e.g., De Groot, De Winter, López-García, Mulder, & Wieringa, 2011; De Winter et al., 2009; Fisher, Pollatsek, & Pradhan, 2006; Roenker, Cissell, Ball, Wadley, & Edwards, 2003; Turpin, Welles, & Price, 2007). Simulators offer features that are not easily attainable through on-the-road training, such as objective performance measurement, manipulation of the environment according to the learning goals, and exposure to errors and hazardous driving conditions in a controlled and repetitive manner without physical risk (Allen, Park, Cook, & Fiorentino, 2007). This study exploits these advantages of driving simulators by temporarily making the driving task more difficult and error-prone. That is, instead of aiming to maximize the learner's performance during training, we deteriorated the vehicle dynamics such that it became more difficult to keep the car in the lane at a given speed.

Previous research on motor-skill and verbal learning concurs that degraded conditions during training, such as variations in task conditions or faster than real-time speeds, can support longer-term learning (Hone & Morrison, 1997; Jarmasz, 2006; Nusseck, Teufel, Nieuwenhuizen, & Bülthoff, 2008; Schmidt & Bjork, 1992; Stefanidis et al., 2007). The exact mechanisms are still to be unraveled, but this phenomenon might be explained from an informational perspective (Lintern, 1991). As Lintern explained, the practice task does not have to be identical to the retention task; rather than physical similarity, critical perceptual similarities form the basis for the transfer of a skill from one task to another. Difficult task conditions may present the learner with increased opportunities for learning such critical invariants. Guadagnoli and Lee (2004) further explained that increased task difficulty during practice provides the learner with an increased information-potential for learning, although care must be taken to not mentally overload the learner.

Research also suggests that the active promotion of errors during training is effective for skill learning (Keith & Frese, 2008; Milot, Marchal-Crespo, Green, Cramer, & Reinkensmeyer, 2010). By making an error, the learner is presented with feedback about the limits of tolerable behavior. When an error occurs, the task is temporarily interrupted and the learner can reflect on why this error occurred, facilitating storage in long-term memory (Frese et al., 1988; Ivancic & Hesketh, 2000). Deteriorated task conditions and the promotion of errors during training may also have positive effects on driving style. Self-induced errors may stimulate meta-cognitive skills and emotional control, and prevent overconfidence (Hogarth, Gibbs, McKenzie, & Marquis, 1991; Ivancic & Hesketh, 2000). Furthermore, through the speed-accuracy trade-off mechanism, drivers are likely to slow down when accurate performance cannot be maintained (Zhai, Accot, & Woltjer, 2004).

The focus in this study was on lane-keeping, an essential driving task that all prospective drivers have to learn. Lane-keeping performance is often used as a proxy variable for road safety (e.g., Brookhuis & De Waard, 1993) and crash statistics reveal that loss-of-control crashes—either due to deficient driving skills or due to deficient driving style—are a concern amongst young drivers (Laapotti & Keskinen, 2004). In this study, participants without a driving license practiced the task of keeping the car near the center of the right lane on a rural road in a driving simulator. Three groups were created. One group trained with a 50% reduction in the normal tire-road friction coefficient, corresponding to a grip level in heavy rain. This grip reduction implies that the maximum speed in a curve of a given radius is decreased by 29%, and that the minimum braking distance from a given speed is doubled. It may be noted that driver training with such low grip tires is unfeasible and illegal in real traffic, but it is possible in a driving simulator or on a closed practice area using ‘‘Skid Car’’ simulation equipment mounted under a car, lifting the car and thereby reducing the grip of the tires (Gregersen, 1996; Jones, 1995). The second group trained with a normal friction coefficient. The third group trained with a 100% increase in the normal tire-road friction coefficient, which corresponds to racing-car tires without tread pattern. This grip level means that the maximum speed in a curve is increased with 41%, and that the minimal braking distance is halved. After the practice sessions, all groups performed an immediate retention session with normal grip. In order to assess a longer term learning effect, a delayed retention session was performed on the next day.

It is noted here that the low-grip condition did not resemble skid-control training, which was previously demonstrated to be ineffective in improving driving style as it increased driver confidence (e.g., Gregersen, 1996; Katila et al., 1996). The car dynamics in the present study were such that once the critical grip level was exceeded while cornering, the car began to slide and loss of control was likely. Hence, the learners were intrinsically motivated to not exceed the acceleration limits of the car. Moreover, the high-grip condition did not resemble a racing game, which was previously found to inflate risk-taking (cf. Fisher, Kubitzki, Guter, & Dieter, 2007). The task instructions in this study focused on lane-keeping accuracy, and participants were not encouraged to drive fast.

Table 1. Experimental hypotheses on the relationship between lane-keeping error and speed. Results are compared with the normal-grip group.

		Lane-keeping error	Speed
Practice	Low grip (LG)	Higher	Lower
	High grip (HG)	Lower	Higher
Retention	Low grip (LG)	Lower	Lower
	High grip (HG)	Higher	Higher

The experimental hypotheses are illustrated in Table 1. We expected that during practice the low-grip group would show higher lane-keeping error and lower speed than the normal-grip group because of the higher task difficulty. We also expected that during retention, that is, when being tested with normal grip level, the low-grip group would have lower lane-keeping error and lower speed than the normal-grip group. As shown in Table 1, we expected the opposite results for the high-grip group. Additionally, we measured workload and confidence to gain a more in-depth understanding of how task conditions during practice and retention tests influence driving perception.

6.2 Method

6.2.1 Participants

Sixty-three people without a driver’s license were recruited from the TU Delft student community. The participants completed an intake questionnaire with the following items: 1) gender (*male*/*female*); 2) age; 3) “Did you take driving lesson(s) on the road?” (*yes*/*no*); 4) “Have you ever driven in a simulator before?” (*yes*/*no*); 5) “Do you play video games for at least 1 hour per week?” (*yes*/*no*); and 6) “I have good steering skills, for example in cycling or computer games”, on a 10-point scale from 1 (*completely disagree*) to 10 (*completely agree*). Of the 63 participants, 21 were female, 25 had started taking driving lessons on the road, 5 had driven in a simulator before, and 26 reported playing video games for more than one hour per week. The mean age was 22.4 years (SD = 3.5), and the mean self-reported steering skill was 6.2 (SD = 2.0). Each participant was compensated with 10 euros.

6.2.2 Apparatus

The driving simulator (Figure 1) was fixed-base and provided a realistic simulation with a 180 degree field of view and surround sound. The simulated car was a middle-class vehicle with a mass of 1,265 kg and a top speed of approximately 180 km/h. The pedals, steering wheel, ignition key and seat resembled those of an actual car, and gear changing was automatic. The virtual world was visually projected by means of three LCD projectors (front projector NEC VT676, brightness 2,100 ANSI lumens, contrast ratio 400:1, resolution 1,024 x 768 pixels; side projectors NEC VT470, brightness 2,000 ANSI lumens, contrast ratio 400:1, resolution 800 x 600 pixels), and the dashboard, interior, and mirrors were integrated into the projected image. The sound in the simulator cab consisted of wind, engine, and tire noises. Tire-squeal was audible only when the tires approached their traction limit during braking; no tire-squeal sound was produced when the tires were approaching their lateral limits. To make the steering wheel force-feel characteristics independent of the grip level of the

Figure 1. The driving simulator that was used in the experiment. The dashboard shows the speed and engine rpm. Gear changing was automated.

tires, the steering wheel system was modified from an active force-feedback system to a passive mass-spring-damper system, otherwise the guidance properties of force feedback (e.g., Winstein, Pohl, & Lewthwaite, 1994) and the physical capabilities of the participant could have interacted with the independent variables and confounded the results. The steering sensitivity was calibrated to correspond to the on-centre characteristics of cars on the road (Katzourakis, De Winter, De Groot, & Happee, 2012). A number of experienced drivers tested the simulator and did not report anything unusual after driving the simulator with the passive steering system.

6.2.3 Experimental groups

There were three experimental groups: Low grip (LG), $n = 22$; Normal grip (NG), $n = 21$; and High grip (HG), $n = 20$. The assignment of each participant to one of the three groups was determined by the Taves (1974) procedure in order to minimize the differences between the groups in terms of four selected variables: (1) gender, (2) age (<21 years vs. >21 years), (3) self-reported steering skill (<5 vs. >5 on the 10-point scale), and (4) driving lessons on the road (yes vs. no). These four variables are known to be predictive of steering performance and driving speed in a driving simulator (Cantin, Lavallière, Simoneau, & Teasdale, 2009; De Groot, De Winter, López-García et al., 2011; De Winter et al., 2006; De Winter, Wieringa, Kuipers, Mulder, & Mulder, 2007; Petzoldt, Bär, & Krems, 2009). Simulation studies have shown that minimization provides better balanced groups than conventional randomization (Scott, McPherson, Ramsay, & Campbell, 2002). LG practiced with a friction coefficient of 0.45, NG practiced with a friction coefficient of 0.90, and HG practiced with a friction coefficient of 1.80. The front and rear tire friction coefficients, and the static and dynamic friction coefficients were equal.

6.2.4 Procedure and tasks

After recruitment, an email was sent to the participants containing the time and location of the experiment, as well as a protocol explaining that the task goal was to drive as accurately as possible near the centre of the right lane. All participants provided written informed consent.

Participants first performed an 8 minute baseline session in the simulator to practice the auditory reaction time task (described below). Six 8 minute driving sessions were then completed: four Practice sessions, an Immediate Retention session, and a Delayed Retention session on the following day. In the retention sessions, all participants were provided with the normal grip tires. All driving sessions took place on a two-lane 7.5 km lap in a country environment without intersections or other vehicles. The lane width was 5 m. The lap consisted of 25 curves of varying angles and radii. The road contained a tunnel and two 4 m hills. The road surface was uniform and flat. Previous research on this simulator showed that performance on lane-keeping and steering tasks predicted the chance of passing the Dutch driving license test (De Winter et al., 2009).

Before each driving session, a series of written instructions were projected in the simulator, explaining how to perform the reaction time task and that the aim of each session was to drive as well as possible in the centre of the right lane. Prior to Immediate Retention, the participants were shown the following text: “Attention! The vehicle and its behavior could be different from the previous sessions”.

During both the baseline session and the five driving sessions, an auditory reaction time task had to be performed. The task was to react as quickly as possible to a 0.1 s beeping sound produced at a random time interval between 4 and 8 s. The reaction time was measured from the moment the beep occurred until the moment the participant pressed the horn (central piece on the steering wheel). After pressing the horn, a second beep with lower tone was produced as a confirmation. No confirmation sound was produced when the participant did not react within 2 s. After each of the six sessions, participants were asked to step out of the simulator to complete the NASA Task Load Index (TLX; Hart & Staveland, 1988), the Rating Scale Mental Effort (Zijlstra, 1993), and our own confidence questionnaire.

6.2.5 Dependent measures

The following dependent measures were calculated for each practice and retention session.

<u>Lane-keeping</u>

Number of departures. This measure counted the number of times that the car crossed the road boundaries. Road departures were typically the consequence of improper lane-keeping behavior or loss of control because the curve was approached too fast. When a road departure occurred, the participant’s car was automatically put back in the centre of the right lane with zero speed and the engine switched off.

RMSE overall (m). The Root Mean Squared Error (RMSE) of the distance between the centre of the vehicle and the centre of the right lane describes how accurately the driver kept the vehicle near the lane centre.

RMSE curves (m). The RMSE of the distance between the center of the vehicle and the center of the right lane during 90 degree curves with a road center-line radius of 15 or 20 m. Only the first eight curves (five right-hand and three left-hand curves) were included in the analysis, because even when driving slowly, drivers could easily complete this part of the driving course. Therefore, the same curves were used in the analysis for all participants (the same eight curves were used for mean LP curves and mean speed curves, see below). The RMSE was calculated for each curve separately and then averaged over all eight curves in order to obtain the RMSE curves measure. Note that previous experiments in the driving simulator showed that the participants' behavior in these sharp curves is a sensitive measure, while the mild curves were relatively unaffected by the experimental conditions (De Groot, De Winter, Mulder, & Wieringa, 2011b).

Mean LP curves (m). The mean distance of the centre of the vehicle to the centre of the right lane during the 90 degree curves provided an indication of the line taken through the curves. A positive LP means that the participant drove to the left of the lane centre; a negative LP means that the participant drove to the right of the lane centre. Mean LP left-hand curves and mean LP right-hand curves were calculated separately.

<u>Speed</u>

Mean speed overall (m/s). The mean speed of the simulated vehicle. A higher speed is assumed to be indicative of a poorer driving style and increased risk-taking, and is associated with an increased risk of accidents (Aarts & Van Schagen, 2006; Elvik, Christensen, & Amundsen, 2004). Mean speed has been used as a dependent variable in many previous driving simulator studies (e.g., Gelau, Sirek, & Dahmen-Zimmer, 2011; Matthews et al., 1998; Reimer, Mehler, Coughlin, Roy, & Dusek, 2011; Santos, Merat, Mouta, Brookhuis, & De Waard, 2005). Although drivers in a low-cost simulator usually drive considerably faster than they would do in a real car, and absolute validity is therefore low, speed and speeding in a simulator is a valid measure as far as relative comparisons are concerned (Bédard, Parkkari, Weaver, Riendeau, & Dahlquist, 2010; De Groot, De Winter, Mulder, & Wieringa, 2011b; Godley, Triggs, & Fildes, 2002; Lee, Lee, Cameron, & Li-Tsang, 2003; Shechtman, Classen, Awadzi, & Mann, 2009).

Mean speed curves (m/s). The mean speed of the simulated vehicle during the 90 degree curves.

<u>Workload</u>

Mean RT (s). The mean reaction time on the auditory reaction time task, representing the participant's mental workload. Reaction times less than 0.1 s (anticipatory responses) and reaction times exceeding 2 s (failed responses) were excluded.

TLX (%). The NASA Task Load Index (TLX) provided an indication of the workload on the following six dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration. A total score from 0% to 100% was calculated by summing the six items. The NASA TLX is a widely used scale designed to obtain workload estimates. It was initially designed for aviation, but is now used in many other applications, including driving in an estimated 9% of the studies using the NASA TLX (Hart, 2006).

Effort. The Rating Scale Mental Effort provided an indication of the participant's level of effort expenditure (Zijlstra, 1993). This rating scale was presented on A4 paper as a 150 mm vertical bar with anchors at nine points, from 2 mm (*absolutely no effort*) to 112 mm (*extreme effort*). The instructions on the form stated: "Indicate by marking on the vertical axis below how much effort it took you to complete the task you have just finished".

<u>Confidence</u>

Confidence (%).The participant's confidence was assessed using our confidence questionnaire, which contained the following six statements: (1) ''I understood how to negotiate the driving situations presented in the simulation'', (2) ‘‘Keeping the car in the center of the lane was easy’’, (3) ‘‘I performed well on keeping the car in the center of the lane’’, (4) ‘‘I think I performed better than the average participant in keeping the car accurately in the center of the right lane’’, (5) ‘‘I had a feeling of risk during driving’’, and (6) ‘‘I feel confident to drive in similar conditions in the real world’’. These items were inspired from previous questionnaires about drivers’ confidence (De Craen, 2010; Ivancic & Hesketh, 2000; Wells, Tong, Sexton, Grayson, & Jones, 2008) and adapted towards the present simulator-based lane-keeping task. Reactions to the statements could be given by marking a cross on a 21 tick horizontal bar identical to those used in the NASA TLX, with anchors on the left (strongly disagree) and right sides (strongly agree). A total confidence score was calculated on a range from 0% to 100% by averaging the six items (the fifth item was reversed).

6.2.6 Statistical analyses

For each dependent measure, and for Practice 1–4, Immediate Retention, and Delayed Retention, three comparisons were made: LG vs. NG, LG vs. HG, and NG vs. HG. The comparisons were performed with a point-biserial correlation coefficient, controlling for initial aptitude. The dependent measures were rank-transformed (Conover & Iman, 1981) for higher robustness and to cope with the skewed distribution of some of the variables, such as the number of departures.

The initial aptitude was calculated from the following information, acquired prior to commencing the first practice session: the six items from the intake questionnaire and the mean reaction time and effort of the baseline session. The matrix of these variables (63 participants × 8 variables) was reduced to one score, representing initial aptitude, by taking the first principal component based on the correlation matrix. The mean initial aptitude scores were -0.12 (SD = 1.07), 0.03 (SD = 0.90), and 0.10 (SD = 1.05) for LG, NG, and HG, respectively. These means were not significantly different as determined with a one-way analysis of variance (F = 0.270, p = .764).

Additionally, comparisons between sessions were carried out with a paired t test. To assess changes during practice, Practice 1 and Practice 4 were compared. To assess changes between practice and retention, Practice 4 and Delayed Retention were compared. Finally, we compared Immediate Retention with Delayed Retention in order to assess the overnight effects.

6.3 Results

Table 2 shows the correlation matrix and the Practice 3 to Practice 4 reliabilities of the dependent measures. The reliabilities were generally greater than .8. A relatively low reliability (.53) was found for the number of departures, which can be explained by the fact that road departures were occasional events, in contrast to the other dependent measures that were all based on continuous variables.

The Mean LP right-hand curves also had a low reliability (.67), which was likely caused by range restriction. In left-hand curves it was possible to cut the corner because the participants had room to use the left lane. In right-hand curves, on the other hand, cutting the corner was less possible because the car would have to be driven onto the verge. Thus, the variability of trajectories in right-hand curves was smaller compared to left-hand curves.

The correlation matrix (Table 2) reveals that the measures describing lane-keeping error (RMSE, RMSE curves, and Mean LP left-hand curves) had high correlations (>.85). This can be explained by the large lane-center error in left-hand

Table 2. Correlation matrix for the dependent measures, with test-retest reliabilities on the diagonal (N = 63).

	1	2	3	4	5	6	7	8	9	10	11	12
1. Number of departures	.53											
2. RMSE (m)	.70	.92										
3. RMSE curves (m)	.72	.92	.80									
4. Mean LP left curves (m)	.64	.86	.89	.73								
5. Mean LP right curves (m)	.16	.24	.08	.05	.67							
6. Mean RT (s)	.33	.46	.44	.23	.14	.94						
7. TLX (%)	.25	.33	.30	.30	.25	.17	.91					
8. Effort	.28	.31	.27	.18	.20	.26	.66	.95				
9. Mean speed (m/s)	.28	−.15	−.16	−.12	−.06	−.33	−.17	−.05	.89			
10. Mean speed curves (m/s)	.49	.31	.34	.38	−.18	.03	.05	.11	.50	.70		
11. Confidence (%)	−.45	−.35	−.39	−.32	−.12	−.13	−.41	−.30	.09	−.18	.87	
12. Initial aptitude	−.21	−.45	−.48	−.44	−.20	−.28	−.36	−.23	.45	.07	.39	

The correlations were determined by first averaging the measures across the six driving sessions. Group effects were eliminated by subtracting the group mean. Correlations of magnitude greater than or equal to .25 are significant, p < .05. The test-retest reliabilities were calculated using the correlation between Practice 3 and 4. The initial aptitude represents the covariate used in the group comparisons.

curves contributing to RMSE. Because the differences between the three groups were similar for RMSE, RMSE curves, and Mean LP left-hand curves, below we only report the results for Mean LP. The TLX and Effort measures were substantially correlated as well (.66). For the sake of simplicity, and because the differences between the three groups were similar for both measures, we only present the results of the TLX. Cronbach's alpha, calculated for each of the six sessions separately, was on average .86 for our confidence questionnaire and .80 for the NASA TLX. All group means of the dependent measures are shown in Figs. 2–8. The corresponding p values and effect sizes are provided in Tables 3 and 4.

Lane-keeping

Number of departures. Figure 2 shows the number of departures. There was a significant performance improvement from Practice 1 to Practice 4 for all three groups. HG had significantly less road departures than the other two groups during practice. From Practice 4 to Immediate Retention, the number of departures significantly increased for HG, whereas it significantly decreased for LG. During Immediate Retention, LG had the lowest number of departures; the difference between LG and HG was statistically significant. From Immediate to Delayed Retention, the number of departures by HG significantly decreased. There were no significant differences between the three groups during Delayed Retention.

Mean LP curves. Figure 3 shows the Mean LP in left- and right-hand curves. The Mean LP is higher for left-hand curves than for right-hand curves, which is due to the above-mentioned phenomenon of corner-cutting. NG and HG had significantly lower Mean LP left curves during Practice 4 than during Practice 1, indicating that they learned to drive closer to the lane centre during the practice sessions. During Immediate Retention, HG cut the left-hand corners significantly less (i.e., drove closer to the lane centre) than the other two groups. In right-hand curves, LG cut the corners more than the other two groups during Immediate Retention, and also more than during Practice 4. The difference between LG and HG in right-hand curves was still significant the following day, that is, during the Delayed Retention session. Figure 4 illustrates the vehicle paths of LG and HG during Immediate Retention, showing that LG drove closer to the inside of the curve than HG.

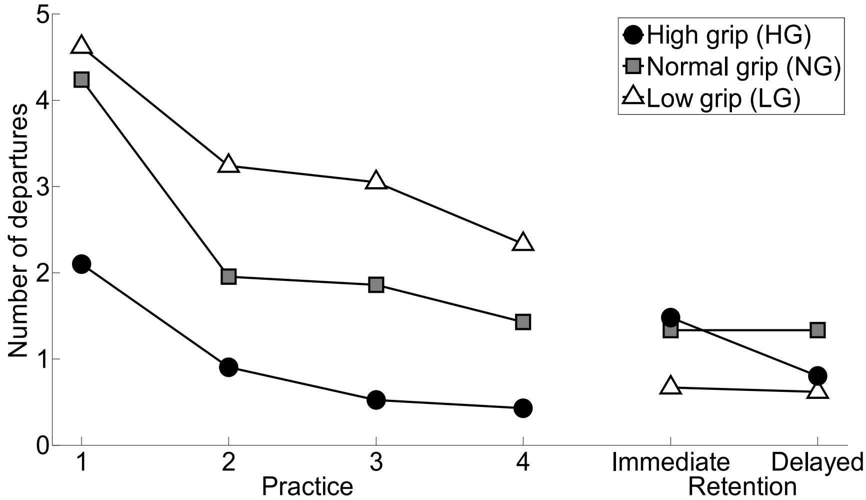

Figure 2. Group averages of the number of road departures during the practice and retention sessions.

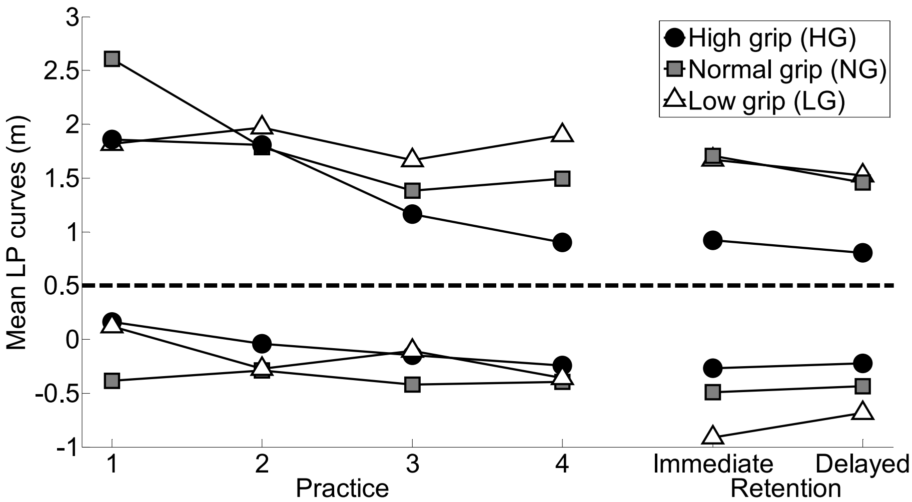

Figure 3. Group averages of the mean lateral position (LP) in curves during the practice and retention sessions. Above the dashed line, mean LP in left curves; below the dashed line, mean LP in right curves.

<u>Speed</u>

Mean speed overall. Figure 5 (top) shows the overall mean speed per session. The mean speed did not change during practice for any of the three groups. During practice, LG drove slower than the other two groups, whereas NG and HG did not differ significantly from each other. From Practice 4 to Immediate Retention and from Immediate to Delayed Retention, LG significantly increased its mean speed. During Immediate Retention, LG still drove significantly slower than NG, but during Delayed Retention there was no significant difference between the three groups.

Mean speed curves. The trends in the group differences for the mean speed in curves (Figure 5, bottom) were similar to those of the overall mean speed. During the practice sessions, LG drove significantly slower than NG and HG. Consistent with the hypothesis, LG drove significantly slower than NG during Immediate and Delayed Retention. HG reduced speed from Practice 4 to Immediate Retention, while LG increased speed from Immediate to Delayed Retention.

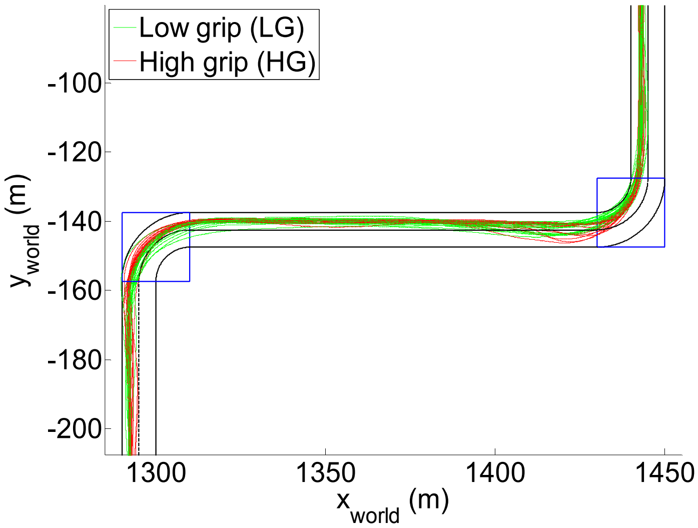

Figure 4. Paths of the centre of the participants' car in the first two 90 degree curves of the Immediate Retention session. Road departures were excluded from this visualization. The squares indicate the regions over which curve measures were calculated. To make the figure clearer, only the paths of the low-grip (LG) and high-grip (HG) groups are shown. It can be seen that LG tended to hug the inside of the curves more than did HG.

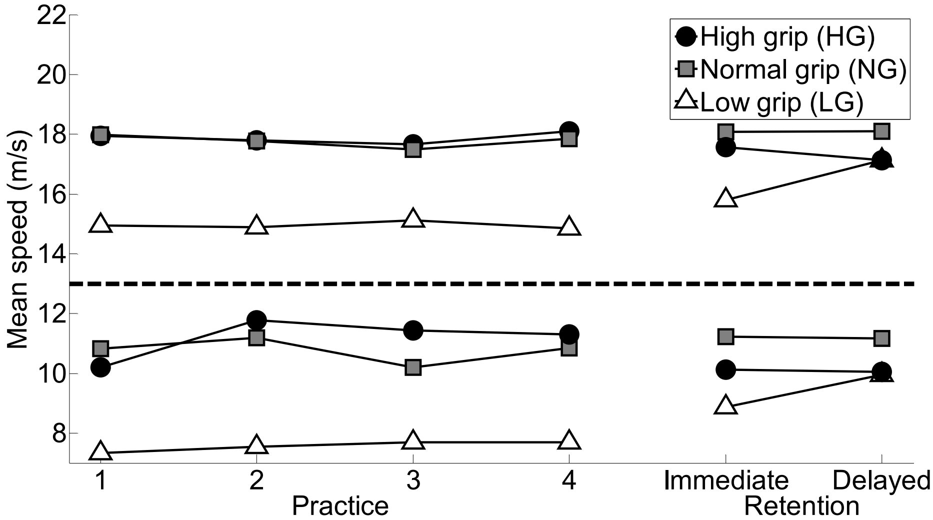

Figure 5. Group averages of the mean speed during the practice and retention sessions. Above the dashed line, overall mean speed; below the dashed line, curve mean speed.

<u>Workload</u>

Mean RT. Figure 6 shows the results for Mean RT. During practice, the Mean RT of the three groups was equal. There was a clear learning effect; all groups significantly improved from Practice 1 to Practice 4. Between Practice 4 and Immediate Retention, LG reduced mean RT, HG increased mean RT, and NG remained similar. During Immediate Retention, HG had the highest reaction time and LG the lowest reaction time; the difference between LG and HG approached significance (p = .057). Mean RT significantly decreased from Immediate Retention to Delayed Retention for all three groups.

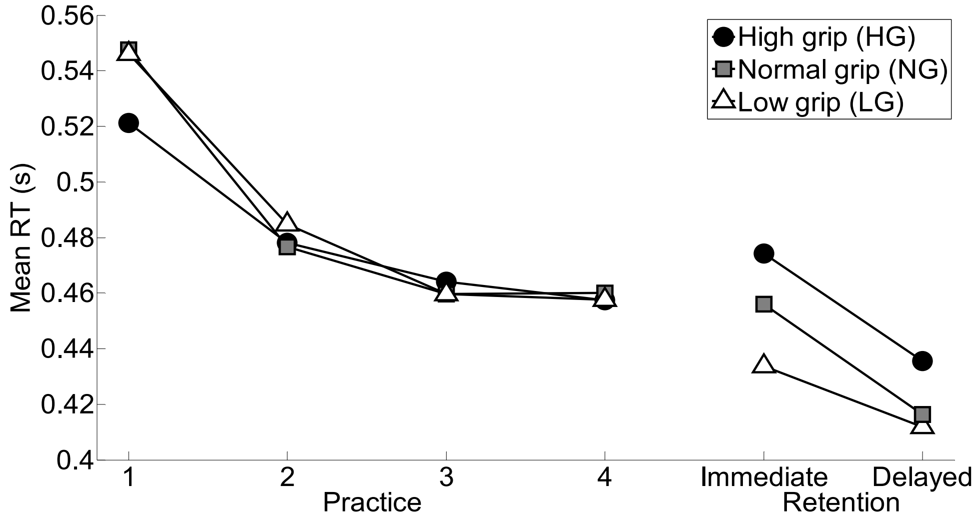

Figure 6. Group averages of the mean reaction time (RT) during the practice and retention sessions.

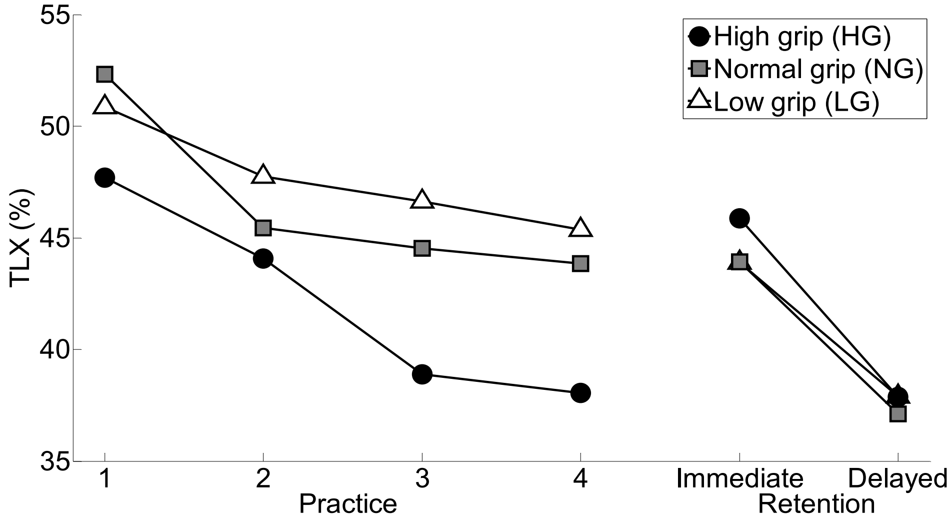

Figure 7. Group averages of the NASA TLX score during the practice and retention sessions.

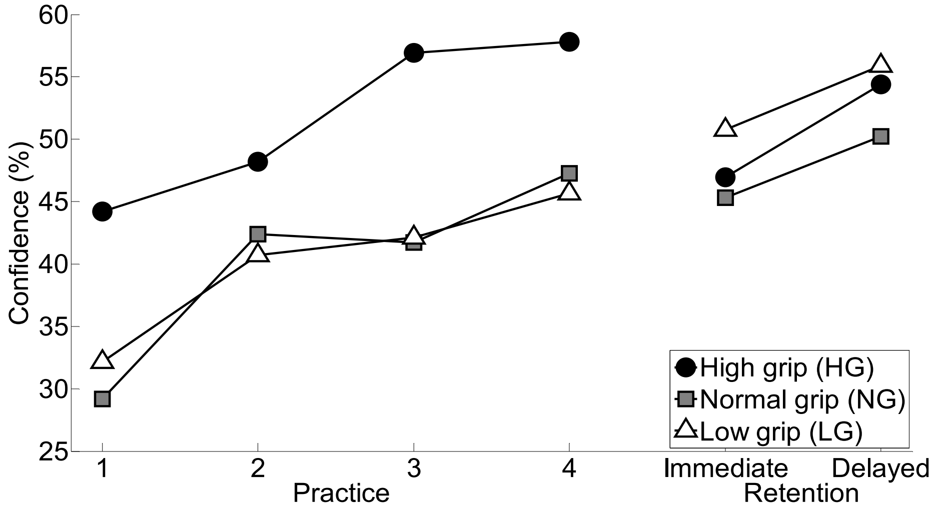

Figure 8. Group averages of the confidence score during the practice and retention sessions.

TLX. Figure 7 shows the responses to the NASA TLX. No group differences were found during practice, but there was a significant reduction of the TLX score from Practice 1 to Practice 4 for all groups. HG showed an increase of the TLX score from Practice 4 to Immediate Retention. As with Mean RT, there was a reduction between Immediate and Delayed Retention. No group differences were found during Immediate and Delayed Retention.

Confidence

Figure 8 shows the confidence questionnaire scores. As hypothesized, HG reported higher confidence during practice than the other two groups. However, confidence scores were similar for LG and NG during practice. All groups showed a strong increase in confidence between Practice 1 and Practice 4. Between Practice 4 and Immediate Retention, HG showed a significant decrease in confidence, whereas LG showed an increase in confidence. All groups significantly increased confidence

Table 3. Group comparisons for the dependent measures.

		Group comparison		
Measure	Session	LG vs. HG	LG vs. NG	NG vs. HG
Number of departures	Practice	**.000 (1.57)**	.192 (0.42)	**.001 (1.16)**
	Imm. Retention	**.019 (−0.81)**	.298 (−0.34)	.318 (−0.33)
	Del. Retention	.850 (−0.06)	.130 (−0.49)	.196 (0.44)
Mean LP left curves	Practice	.369 (0.29)	.286 (−0.34)	.087 (0.57)
	Imm. Retention	**.003 (1.07)**	.717 (−0.12)	**.001 (1.29)**
	Del. Retention	.145 (0.50)	.951 (−0.02)	.138 (0.51)
Mean LP right curves	Practice	.178 (−0.44)	.410 (0.26)	**.018 (−0.80)**
	Imm. Retention	**.001 (−1.22)**	**.032 (−0.75)**	.122 (−0.55)
	Del. Retention	**.026 (−0.77)**	.212 (−0.41)	.246 (−0.39)
Mean speed overall	Practice	**.000 (−1.64)**	**.000 (−1.52)**	.676 (−0.14)
	Imm. Retention	.088 (−0.58)	**.028 (−0.73)**	.676 (0.14)
	Del. Retention	.998 (0.00)	.384 (−0.28)	.461 (0.25)
Mean speed curves	Practice	**.000 (−2.71)**	**.000 (−2.97)**	.526 (−0.21)
	Imm. Retention	**.043 (−0.70)**	**.000 (−1.58)**	.083 (0.61)
	Del. Retention	.612 (−0.17)	**.007 (−0.91)**	.090 (0.58)
Mean RT	Practice	.829 (−0.07)	.889 (−0.04)	.896 (−0.04)
	Imm. Retention	.057 (−0.65)	.513 (−0.21)	.284 (−0.36)
	Del. Retention	.233 (−0.40)	.690 (−0.13)	.430 (−0.27)
TLX	Practice	.413 (0.27)	.996 (0.00)	.342 (0.31)
	Imm. Retention	.543 (−0.20)	.944 (0.02)	.387 (−0.29)
	Del. Retention	.985 (0.01)	.946 (0.02)	.926 (−0.03)
Confidence	Practice	**.011 (−0.85)**	.779 (0.09)	**.009 (−0.89)**
	Imm. Retention	.130 (0.51)	.233 (0.39)	.813 (0.08)
	Del. Retention	.669 (0.14)	.120 (0.51)	.361 (−0.31)

The table shows the *p*-values for group comparisons during Practice (four sessions averaged), Immediate Retention, and Delayed Retention. Effect sizes are reported between parentheses. The effect size is Cohen's *d*, that is, the standardized mean difference obtained by converting the point-biserial correlation coefficient using the equation $d = 2r/(1-r^2)$. According to Cohen (1992), *d*s of 0.20, 0.50, and 0.80 can be interpreted as small, medium, and large effects, respectively.

Table 4. Session differences for the three experimental groups.

Measure	Session	Experimental group LG	NG	HG
Number of departures	P1 vs. P4	**.000**	**.000**	**.012**
	P4 vs. Imm. Retention	**.000**	.217	**.000**
	Imm. vs. Del. Retention	.879	.486	**.011**
Mean LP left curves	P1 vs. P4	.354	**.012**	**.003**
	P4 vs. Imm. Retention	.815	.189	.697
	Imm. vs. Del. Retention	**.039**	**.046**	.538
Mean LP right curves	P1 vs. P4	.073	.682	**.020**
	P4 vs. Imm. Retention	**.002**	.790	.547
	Imm. vs. Del. Retention	.088	.894	.718
Mean speed overall	P1 vs. P4	.957	.657	.782
	P4 vs. Imm. Retention	**.000**	.462	**.028**
	Imm. vs. Del. Retention	**.005**	.882	.440
Mean speed curves	P1 vs. P4	.495	.928	.326
	P4 vs. Imm. Retention	**.000**	.376	**.002**
	Imm. vs. Del. Retention	**.000**	.292	.688
Mean RT	P1 vs. P4	**.000**	**.000**	**.001**
	P4 vs. Imm. Retention	**.006**	.320	**.002**
	Imm. vs. Del. Retention	**.003**	**.003**	**.001**
TLX	P1 vs. P4	.092	**.002**	**.017**
	P4 vs. Imm. Retention	.746	.806	**.005**
	Imm. vs. Del. Retention	**.011**	**.006**	**.017**
Confidence	P1 vs. P4	**.001**	**.000**	**.000**
	P4 vs. Imm. Retention	**.028**	.305	**.001**
	Imm. vs. Del. Retention	**.039**	**.016**	**.003**

The table shows *p*-values for within-group changes between Practice 1 and Practice 4, between Practice 4 and Immediate Retention, and between Immediate Retention and Delayed Retention.

in Delayed Retention as compared to Immediate Retention. During Immediate and Delayed Retention no significant group differences were found.

6.4 Discussion

This simulator-based study investigated the effects of the tire-road friction coefficient on learning a lane-keeping task. We hypothesized that practicing with low-grip tires would result in lower lane-keeping error and lower speeds during normal-grip retention tests than practicing with normal-grip tires. We further hypothesized that practicing with high-grip tires would have the opposite effects. Table 5 summarizes the results of the experiment. As hypothesized, the low-grip group (LG) drove significantly slower than the normal-grip group (NG) during the practice and retention sessions. With respect to lane-keeping error, mixed results were obtained: LG had a lower number of road departures than HG during the immediate retention session, but HG drove closer to the lane center than the other two groups.

Low- and high-grip practice did not result in opposite effects. During both practice and retention, HG drove with similar speeds to NG even though the participants were provided with a very high grip level. This reveals that HG participants did not seek to extend their limits in terms of speed, but focused on lane-keeping performance instead. Consistent with the hypothesis, HG had less road departures, lower lane-keeping error, and reported higher confidence than NG during practice. Contrary to expectations, HG did not show higher lane-keeping error than NG in the retention sessions. In fact, during the immediate retention session, HG adhered significantly better to the lane center in curves than NG. Although the confidence of HG was clearly elevated during practice, this higher confidence dissipated when confronted with the normal-grip tires after Practice 4 in the immediate retention session. This transfer to a lower grip level also caused an increase in the number of road departures and a reduction in speed as compared to the final practice session. In conclusion, practicing with high-grip tires did not have an adverse effect on speed and it helped to improve task performance according to the task goal, that is, to drive accurately near the lane center.

An interesting phenomenon was that during the immediate retention session, HG had more road departures than LG, but adhered to the lane center in curves more accurately than the other two groups. Although, strictly speaking, a lower lane-center error signals superior lane-keeping in accordance with the task instructions, it is perhaps not the safest way to drive. NG and LG appeared to cut corners, taking a trajectory that resembled a racing line, which increases the effective cornering radius compared to a constant-radius turn (Beckman, 1991). Hence, for a given driving speed, adopting a racing line reduces lateral accelerations and thus can be considered a safe way to drive. Of course, cutting corners by driving in the adjacent lane can only be recommended when no other traffic is present, as was the case in the present experiment.

A remarkable result was that the workload measures (mean RT & NASA TLX) were not significantly different between the three groups during practice. In other words, even though we manipulated the vehicle dynamics considerably, participant workload was unaffected. The lack of group differences during practice can be explained by the drivers' inclination to compensate their driving behavior in order to keep their mental workload within set boundaries in a homeostatic fashion (Fuller, 2005; for review and discussion see De Winter & Happee, 2012; Lewis-Evans, De Waard, Jolij, & Brookhuis, 2012; Lewis-Evans & Rothengatter, 2009). LG compensated for the low-grip conditions by reducing speed, whereas HG used the high-grip car to improve accuracy, that is, to minimize the number of road departures as compared to NG. However, Fuller's model cannot explain why workload decreased from Practice 1 to Practice 4 while participants did not compensate for this by increasing speed during practice. These results reveal that drivers' behavioral compensation is an intricate mechanism. As explained by Elvik (2004), the amount of behavioral compensation depends on multiple factors, including, for example, how easily a driver notices a change in the driving conditions and whether or not utility can be gained. Differences in workload only appeared during the immediate retention session, during which all groups drove with the normal-grip tires. Comparing Practice 4 to Immediate Retention, Mean RT of LG dropped significantly, no difference in Mean RT was evident for NG, while Mean RT of HG significantly increased. This indicates that HG had a more difficult experience when confronted with the nominal task conditions. For future research we recommend to use other potentially more sensitive workload measures, such as event-related potentials (Brookhuis & De

Waard, 2010; Wu, Liu, & Quinn-Walsh, 2008) or questionnaires that are specifically devoted to mental workload in car driving, such as the Driving Activity Load Index (DALI; Pauzié, 2008; Pauzié & Pachiaudi, 1997).

It is useful to contrast the present research with a driving-simulator study by Ivancic and Hesketh (2000). They evaluated a simulator-based training program, in which learners drove through a virtual environment containing several dangerous situations. An intervention group received explicit feedback on errors (e.g., participants who drove through a red traffic light incurred a fine, indicated by a loud police siren). A control group drove the same scenario, but here failure to use correct strategies did not lead to error feedback. The results showed that training with explicit feedback about errors led to reduced confidence in driving skill and significantly better transfer performance during transfer tests. Our study differs from Ivancic and Hesketh's (2000) study because the independent variable in the latter was feedback after an error, while we manipulated the driving task difficulty and thus implicitly varied the number of errors committed by participants. The feedback in our study after a road departure consisted of stopping the simulation, and then repositioning the car of the participant in the middle of the right lane with zero speed and the engine switched off. This strategy did not place extra emphasis on the error, but it made it clear to the participant that the road departure was undesired behavior and the repositioning gave participants the opportunity to briefly reflect on the error.

What are the implications of this research for simulator-based driver training? Clearly more research needs to be conducted before definite conclusions can be drawn, but this study showed that driver training effectiveness was enhanced by increasing the task difficulty. Classic driver training on the road follows the opposite pattern. That is, driver training is traditionally characterized by practicing under relatively protected conditions with the help of driving instructors (Groeger & Banks, 2007), whereas solo driving is a mentally overwhelming activity for newly licensed drivers (Lee, 2007). As was explained by Schmidt and Wulf (1997), ''it is understandable that well-motivated trainers would want to use whatever means they can to facilitate performance in practice'' (p. 524). Our results suggest that it is effective to increase the task difficulty during practice and allow learner drivers to find the limits of tolerable behavior for themselves by making errors at their own pace. The driving simulator offers a safe environment for this kind of training.

An important limitation of this study is that the retention tests were conducted in a driving simulator rather than in a real car. That is, the present study used a so-called quasi-transfer methodology, which essentially does not provide evidence about the way drivers learn to drive a real car. Although there is ample evidence that shows that driving-simulator measures are predictive of on-the-road performance (e.g., De Winter et al., 2009; Kraft, Amick, Barth, French, & Lew, 2010; Shechtman et al., 2009), only a few studies have previously investigated whether skills learned in a driving simulator transfer to the road (Strayer & Drews, 2003; Uhr, Felix, Williams, & Krueger, 2003). In the field of aviation, studies on the transfer of training are much more common (Jacobs, Prince, Hays, & Salas, 1990) and it appears that the quasi-transfer of training methodology is valid in many, but not all, cases (Taylor, Lintern, & Koonce, 1993). One major difference between driving in a simulator and driving in a real car is the perception of haptic and vestibular information (Kemeny & Panerai, 2003). The fixed-base simulator as configured in the present study did not provide haptic and vestibular motion cues, but provided only visual and auditory information about the driving task. In real-world car driving, the human haptic and vestibular sensory systems are useful for detecting acceleration, particularly at high frequencies. In contrast, the visual system is used for detecting relatively slow changes in position and attitude (Brown, Cardullo, & Sinacori, 1989). Haptic and vestibular cues may be particularly relevant during high speed cornering or when driving on the verge of stability (De Winter, Dodou, & Mulder, 2012). It would be of great value to the driver training community if transfer of training studies were conducted with the aim of establishing which aspects of driving skill and driving style learned in the simulator can be generalized to the operational environment.

Table 5. Summary of the experimental results. Results in bold are in agreement with the hypotheses (Table 1).

		Lane-keeping error	Speed
Practice	Low grip (LG)	n.s.	**Lower**
	High grip (HG)	**Lower**	n.s.
Retention	Low grip (LG)	Mixed*	**Lower**
	High grip (HG)	Mixed*	n.s.

*HG adhered better to the lane centre in curves than did LG and NG. However, LG had significantly less road departures than HG during Immediate Retention.

Chapter 7. On the way to pole position: The effect of tire grip on learning to drive a racecar

Abstract

Racecar drivers could benefit from new training methods for learning to drive fast lap times. Inspired by the learning-from-errors principle, this simulator-based study investigated the effect of the tire-road friction coefficient on the training effectiveness of a car racing task. Three groups of 15 inexperienced racecar drivers (low grip (LG), 66% of normal grip; normal grip (NG); high grip (HG), 150% of normal grip) completed four practice sessions of 10 minutes in a Formula 3 car on an oval track of 800 m. After the practice sessions, two retention sessions followed: a retention session with normal grip in a Formula 3 car and another retention session with a Formula 1 car. The results showed that LG was significantly slower than HG in the first retention session. Furthermore, LG reported a higher confidence and lower frustration than NG and HG after each of the two retention sessions. In conclusion, practicing with low grip, as compared to practicing with normal or high grip, resulted in increased confidence but slower lap times.

De Groot, S., & De Winter, J. C. F. (2011). On the way to pole position: the effect of tire grip on learning to drive a racecar. *Proceedings of the IEEE Conference on Systems, Man and Cybernetics,* Anchorage, AK, 133–138.

7.1 Introduction

The primary goal of a driver during the qualifying session of a car-racing event is to drive the fastest possible lap time. In a typical race format, the fastest lap time of each driver during this session determines the starting order for the ensuing race. Racing drivers, like other athletes, are interested in training methods that prepare them better than their competitors.

Driving simulators are powerful training tools because they provide several advantages over on-road driving, including a) inexpensive training and testing time, b) experimental control of the environment, c) accurate measurements of the vehicle state, and d) a safe environment for the driver. In a simulator, drivers are able to explore the limits of their behavior without risking serious consequences. A simulator thus offers the possibility to learn from errors, something which is much more restricted in reality. There is a growing body of evidence showing that performance in a driving simulator is predictive of performance in real cars and that skills learned in the simulator transfer to new situations (Lee & Lee, 2005; Bédard, Parkkari, Weaver, Riendeau, & Dahlquist, 2010; De Winter, De Groot, Mulder, Wieringa, Dankelman, & Mulder, 2009; Shechtman, Classen, Awadzi, & Mann, 2009; Ivancic & Hesketh, 2000; Roenker, Cissell, Ball, Wadley, & Edwards, 2003; Uhr, Felix, Williams, & Krueger, 2003). In this study we adopt a novel approach to simulator-based driver training, by intentionally degrading the handling characteristics of the car with the aim to learn from errors.

Research in motor and verbal learning has shown that deteriorated practice conditions can have a positive effect on posttraining retention performance. A review of Schmidt and Bjork (1992) showed that random practice is better than blocked practice, a reduction of the feedback frequency is better than more feedback, and varying task conditions are better than constant task conditions, when the aim is to enhance performance during retention tests. These practice conditions reduced task performance during practice but improved the level of performance in the long term and in altered contexts.

The racing task is considerably more complex than the motor and verbal tasks reported by Schmidt and Bjork. Car racing is a continuous task during which the driver controls the vehicle though different control interfaces: a throttle and a brake pedal for longitudinal control, and a steering wheel for lateral control. The dynamics are of second order, meaning that the inputs influence the vehicle’s acceleration. Furthermore, car dynamics are nonlinear when driving near the limit of the tires.

Concerning the training of complex tasks, a study about the training of helicopter flying has shown that pilots training in an agile helicopter performed better when being tested in a sluggish helicopter than vice versa (Nusseck, Teufel, Nieuwenhuizen, & Bülthoff, 2008). In previous work, we conducted an experiment in a passenger car driving simulator where the grip of the car was modified to influence task difficulty (De Groot, Centeno Ricote, & De Winter, 2012). The results showed that the four 8-min practice sessions with low grip resulted in a lower speed during two 8-min retention sessions (of which the second one was administered the following day) and a workload reduction from practice to retention. The high-grip group performed the task of keeping the car near the center of the lane better than the normal-grip and low-grip groups and also had a higher confidence during practice. The higher confidence of the high grip group disappeared in the retention sessions. In summary, previous research has shown that task conditions during practice have an influence on retention performance, and that the level of confidence and workload can quickly change when the task conditions change when transferring from practice to retention.

In the present study, the tire-road friction coefficient was altered. Three groups were compared: low grip (LG) with a maximum tire-road friction coefficient of 1.1, normal grip (NG) with a maximum tire-road friction coefficient of 1.7, and high grip (HG) with a maximum tire-road friction coefficient of 2.6. After the training, all groups drove a retention session in the same Formula 3 (F3) car with normal grip and; after that they drove another retention session with a Formula 1 (F1) car. Table I and Fig. 1 show data of record laps driven by an experienced racecar driver for the three different grip levels in the F3 car, and for the F1 car. With high grip, no braking was necessary and the driver could be aggressive with the throttle pedal and steering wheel without losing control. With normal grip, braking was necessary before the corners and the throttle and steering wheel had to be controlled with more caution than with high grip. The low-grip condition resulted in the lowest cornering speeds; the driver had to actively control the steering, brakes, and throttle for a larger proportion of the lap resulting in the most difficult and error-prone practice conditions.

Errors provide potential for learning, and errors during training may benefit retention performance (Ivancic & Hesketh, 2000). Errors may teach drivers what the limit of grip is and how to approach that limit. When the low-grip group receives more grip during the first retention session the driving task becomes less difficult for them. We expected this change in task difficulty to increase the self-confidence of LG, and accordingly we expected LG to explore the limits of the car further, which should result in faster lap times. The following hypothesis was therefore tested: Participants practicing with low grip drive faster lap times in nominal-condition retention sessions than participants practicing with normal or high grip.

Table 1. Characteristics of record laps driven by an experienced racecar driver in a driving simulator.

	LG	NG	HG	F1
Lap time (s)	22.37	18.23	15.27	13.83
Max. speed (km/h)	181	197	213	266
Min. speed (km/h)	91	118	168	152
Brake pressed (% of time)	24	21	0	14
Full throttle (% of time)	35	69	82	57

These laps were driven in identical task conditions as during the experiment.

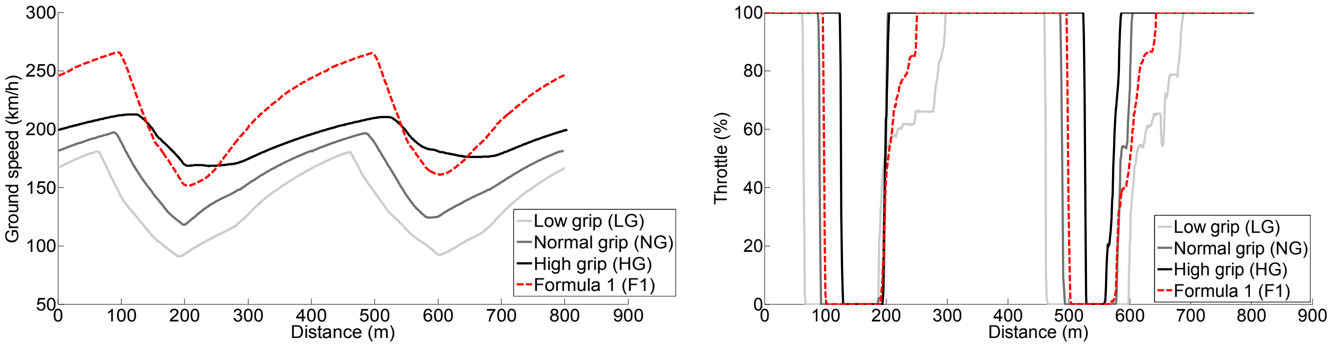

Figure 1. Ground speed (left) and throttle position (right) as a function of distance travelled during the lap, for the four laps reported in Table I. Note that these laps were driven in identical task conditions as the experiment reported later in this paper.

7.2 Method

7.2.1 Apparatus

The racecar simulator consisted of the chassis of a Formula racing car (Figure 2). The steering wheel, throttle system and brake pedal of the original car were used. The visual system consisted of a 52-inch LCD screen (Sony KDL52-5000) positioned 1.3 m in front of the driver's face, resulting in a horizontal and vertical field of view of 27 and 46 degrees, respectively. The active force-feedback system normally used on the simulator was disengaged in order to avoid learning effects caused by haptic guidance (Winstein, Pohl, & Lewthwaite, 1994), as well as to prevent driver fatigue. The virtual world and vehicle dynamics were simulated with the rFactor simulation program (v1.255).

7.2.2 Participants

Three groups of 15 male volunteers, all Delft University students, were created through a randomization process. Young males are a typical target group for evaluating a training method for car racing.

At the start of the experiment, each participant was asked to pick a piece of paper without replacement from a basket containing 45 folded pieces of paper; 15 for group 1 (the low-grip group), 15 for group 2 (the normal-grip group), and 15 for group 3 (the high-grip group). An intake questionnaire was administered resulting in the following data concerning the participants: the average age was 23.0 years (*SD* = 1.4), the average time since licensure was 4.6 years (*SD* = 1.4), the participants played racing games on average 0.6 hours per week (*SD* = 1.7), none of the participants had real-world racing experience, and the average response to the statement "I have good steering skills, for example with cycling, car driving, or computer games" was 7.7 (*SD* = 1.1; anchors: 1 = completely disagree, 10 = completely agree).

7.2.3 Instructions

Participants were provided with written instructions on paper. The goal during each driving session was to drive the fastest possible lap. The participants were clearly informed which sessions were practice sessions and which sessions were retention sessions. They were also informed about the changes in grip level, but no indications

Figure 2. The racing simulator during the experiment.

were given about the magnitude of the differences in friction coefficient between groups. Also, no indications about the record lap times or the fastest times of other participants were provided. As an extra motivator, participants were invited before the experiment to enter a challenge for the fastest lap time in the second retention session. The prize consisted of a ticket for two karting sessions during a driver selection event of the Delft university Formula Student racing team. All participants provided written informed consent and agreed not to talk about their lap times to other participants.

7.2.4 Track

The racing track was an oval track, had a length of 800 m, and consisted of two 180° corners. A short lap ensures the same lap is driven many times in one driving session, which is beneficial to evaluate the learning process.

7.2.5 Procedures

Each participant drove six 10-min sessions: four practice sessions and two retention sessions. Between sessions, a 5-min break was held to give all participants the opportunity to relax and prepare for the next session.

After participants stepped out of the simulator, they were asked to complete the NASA-TLX questionnaire to measure the workload (Hart & Staveland, 1988) and a confidence questionnaire.

7.2.6 Independent variable

The tire-road friction coefficient was the independent variable during this experiment. During the practice sessions all groups drove with a Formula 3 (F3) car (mass = 551 kg, maximum power = 164 kW), with different grip level for the three groups. The maximum friction coefficient (μ) was 1.1, 1.7, and 2.6 for the low-grip, normal-grip and high-grip groups respectively. During the first retention session all participants drove with a Formula 3 car with normal grip level (μ = 1.7). During the second retention session a Formula 1 (F1) car was used (mass = 607 kg, maximum power = 537 kW, μ = 2.4). This was done to confront all participants with a new challenge, a higher task difficulty with greater time pressure, in order to evaluate transfer of learning. The tire wear, effects of tire temperature, and fuel consumption were all disabled in order to guarantee constant grip levels within each session.

7.2.7 Dependent measures

Lap times. Lap times are the most important performance measure for this car racing task. Just as in a real qualifying session, the primary task goal is to drive the fastest lap within the 10-minute session. We also calculated the average lap time per session. Rejected laps were excluded from the analyses.

Rejected laps. Laps during which the wall was touched, the asphalt was left with all four wheels, or laps in which the car drove slower than 30 km/h were automatically excluded from all the analysis and classified as rejected laps.

FullThrottle. The full-throttle percentage is a measure of how much energy was put into the longitudinal movement of the car. This measure was calculated as the time-percentage that the throttle control input was maximal during all valid (non-rejected) laps of the session. FullThrottle can also be seen as a measure of task difficulty; the higher FullThrottle, the easier the driving task. When the throttle is fully opened, the car is limited by the engine power and the driver can focus on the lateral control of the car with the steering wheel to position for the next curve.

Workload. The NASA-TLX questionnaire was employed to measure the workload. The NASA-TLX questionnaire comprises six statements to which participants have to respond on a 21-tick bar.

Confidence. In the confidence questionnaire, participants had to respond to the following three statements (anchors: strongly disagree, strongly agree on a 21-tick bar): 1) I had a feeling of risk during driving, 2) I feel confident to drive in similar conditions in the real world, and 3) I think I had a faster lap time than the average participant in my group.

7.2.8 Statistical analyses

The dependent variables were compared per session for LG vs. NG, LG vs. HG, and NG vs. HG using the independent samples Student's *t* test.

7.3 Results

Lap times. Tables 2 and 3 present the fastest lap times and the average lap times per session, respectively. The fastest and average lap times were significantly different between all groups during the practice sessions with the different grip levels. During Retention 1, with equal grip for all groups, the fastest and average lap times of LG and HG differed significantly from each other. No group differences were found for Retention 2.

Rejected laps. Table 4 shows the rejected laps per session. During all practice sessions, HG had less rejected laps than LG. HG had the least rejected laps, also less than NG, but this difference was only significant in the second and third practice session. In the first lap of Retention 1 (see Figure 3) this effect was reversed: HG had significantly more rejected laps than LG (p = .002) and NG (p = .028). As the Retention 1 session progressed, this effect diminished. No differences were found for the complete Retention1or for the Retention 2 session.

FullThrottle. The results are shown in Table 5 and Figure 4. All groups differed from each other during all practice sessions. LG had the lowest full-throttle percentages, then NG, and HG had the highest percentages. During Retention 1, LG had a significantly lower full-throttle percentage than the other two groups.

Workload. Because the differences between the groups were most pronounced with respect to the frustration element of the NASA-TLX questionnaire, we only took this item into account. The results are shown by Table 6. The NASA-TLX showed that the frustration level was not different between the three groups during the practice sessions. During Retention 1 and Retention 2, the frustration of LG was significantly lower than the frustration of NG and HG.

Confidence. The differences between groups were most pronounced for the third item of the confidence questionnaire. This item showed that during Retention 1 and Retention 2, LG was significantly more confident that they were faster than the average participant in their group than NG and HG. These results can be found in Table 7 and are visually illustrated in Figure 5.

7.3.1 Predictors of fast lap times

As a supplementary analysis, we investigated which variables predicted the fastest lap time during Retention 2, which is the session that was driven with the Formula 1 car. Single-order correlations were calculated after pooling the data of all 45 participants. The fastest lap time correlated significantly with the average lap time (r = .78, p < .001) and with the fastest lap time during Retention 1, which was driven with

Table 2. Group averages (*SD*) of the fastest lap time (s) and *p*-values for group comparisons.

Group	Practice 1	Practice 2	Practice 3	Practice 4	Retention 1	Retention 2
LG	24.24 (0.71)	23.90 (0.57)	23.76 (0.50)	23.76 (0.45)	19.18 (0.30)	14.81 (0.28)
NG	19.39 (0.35)	19.19 (0.28)	19.12 (0.30)	19.09 (0.34)	19.03 (0.27)	14.95 (0.48)
HG	16.19 (0.33)	15.90 (0.21)	15.88 (0.20)	15.80 (0.20)	18.97 (0.19)	14.75 (0.28)
LG vs. NG	.000	.000	.000	.000	.169	.379
LG vs. HG	.000	.000	.000	.000	.032	.526
NG vs. HG	.000	.000	.000	.000	.483	.190

Table 3. Group averages (*SD*) of the average lap time (s) and *p*-values for group comparisons.

Group	Practice 1	Practice 2	Practice 3	Practice 4	Retention 1	Retention 2
LG	25.56 (1.28)	24.60 (0.77)	24.51 (0.65)	24.45 (0.55)	19.76 (0.34)	15.67 (0.45)
NG	20.20 (0.73)	19.83 (0.43)	19.64 (0.36)	19.65 (0.36)	19.54 (0.34)	16.02 (0.70)
HG	16.86 (0.43)	16.32 (0.24)	16.26 (0.29)	16.19 (0.27)	19.42 (0.29)	15.83 (0.59)
LG vs. NG	**.000**	**.000**	**.000**	**.000**	.114	.134
LG vs. HG	**.000**	**.000**	**.000**	**.000**	**.009**	.432
NG vs. HG	**.000**	**.000**	**.000**	**.000**	.300	.456

Table 4. Group averages (*SD*) of the rejected laps (% of total laps) and *p*-values for group comparisons.

Group	Practice 1	Practice 2	Practice 3	Practice 4	Retention 1	Retention 2
LG	58 (17)	45 (18)	47 (14)	45 (16)	33 (18)	29 (8)
NG	45 (23)	37 (22)	35 (20)	34 (17)	33 (14)	27 (9)
HG	33 (17)	24 (13)	19 (10)	23 (13)	41 (18)	28 (14)
LG vs. NG	.114	.331	.071	.104	.909	.513
LG vs. HG	**.001**	**.001**	**.000**	**.000**	.214	.696
NG vs. HG	.107	**.049**	**.012**	.054	.201	.928

Table 5. Group averages (*SD*) of the percentage full throttle per lap and *p*-values for group comparisons.

Group	Practice 1	Practice 2	Practice 3	Practice 4	Retention 1	Retention 2
LG	19 (7)	23 (7)	25 (5)	24 (5)	45 (5)	35 (5)
NG	40 (9)	46 (5)	49 (5)	51 (5)	51 (5)	35 (6)
HG	62 (8)	67 (8)	70 (8)	70 (6)	52 (4)	34 (5)
LG vs. NG	**.000**	**.000**	**.000**	**.000**	**.004**	.694
LG vs. HG	**.000**	**.000**	**.000**	**.000**	**.001**	.386
NG vs. HG	**.000**	**.000**	**.000**	**.000**	.704	.717

the Formula 3 car (r = .62, p < .001). However, the fastest lap time did not correlate significantly with the number of rejected laps (r = .06, p = .690) and the full-throttle percentage (r = −.20, p = .188), nor with the confidence (r = −.05, p = .746; see also Figure 6), frustration (r = −.08, p = .599), age (r = .13, p = .396), years of having a driving license (r = −.01, p = .936), number of racing game hours per week (r = −.18, p = .242), and self-reported steering skills (r = −.05, p = .746). These results indicate that the best predictor of fast lap times is previous performance and not self-reported behaviors or skills.

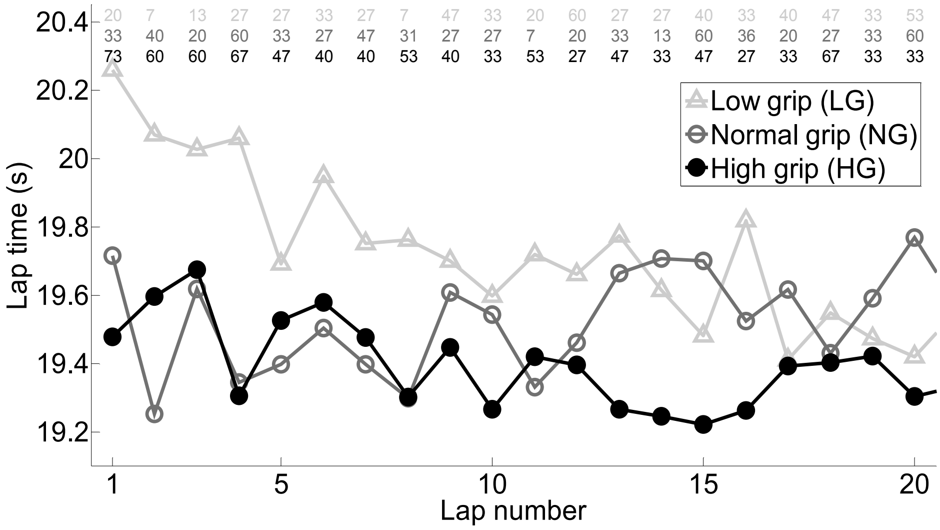

Figure 3. Group averages of the lap time as a function of lap number in the Retention 1 session. Above the graph the percentage of rejected laps is shown per group.

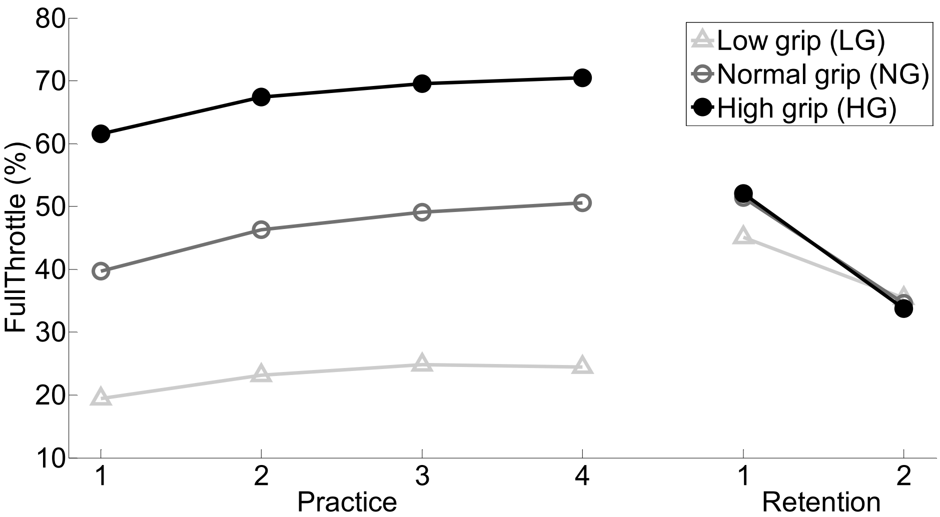

Figure 4. Group averages of the percentage full throttle.

Table 6. Group averages (*SD*) of Frustration item ("How insecure, discouraged, irritated, stressed, and annoyed were you?" percentage from low to high) of the NASA-TLX and *p*-values for group comparisons.

Group	Practice 1	Practice 2	Practice 3	Practice 4	Retention 1	Retention 2
LG	48 (31)	31 (13)	44 (21)	41 (23)	16 (11)	24 (21)
NG	50 (18)	52 (20)	52 (24)	52 (21)	52 (23)	47 (24)
HG	54 (23)	52 (21)	45 (23)	53 (22)	57 (21)	55 (21)
LG vs. NG	.832	**.002**	.342	.157	**.000**	**.010**
LG vs. HG	.577	**.002**	.903	.157	**.000**	**.000**
NG vs. HG	.631	.930	.426	.966	.571	.337

Table 7. Group averages (*SD*) of Statement 3 ("I think I had a faster lap time than the average participant in my group." percentage from strongly disagree to strongly agree) of the confidence questionnaire and *p*-values for group comparisons.

Group	Practice 1	Practice 2	Practice 3	Practice 4	Retention 1	Retention 2
LG	28 (16)	42 (17)	49 (22)	49 (16)	56 (19)	61 (17)
NG	41 (21)	43 (24)	43 (21)	46 (24)	39 (22)	45 (24)
HG	41 (19)	43 (18)	52 (17)	50 (17)	33 (18)	41 (15)
LG vs. NG	.055	.862	.503	.688	**.027**	**.047**
LG vs. HG	.052	.800	.617	.867	**.002**	**.003**
NG vs. HG	.927	.966	.212	.600	.475	.656

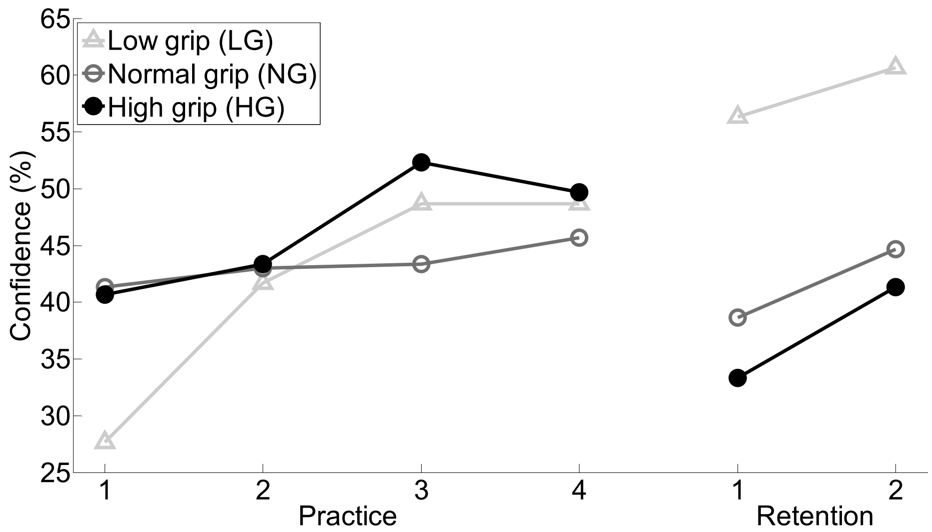

Figure 5. Group averages for Statement 3 of the Confidence questionnaire ("I think I had a faster lap time than the average participant in my group").

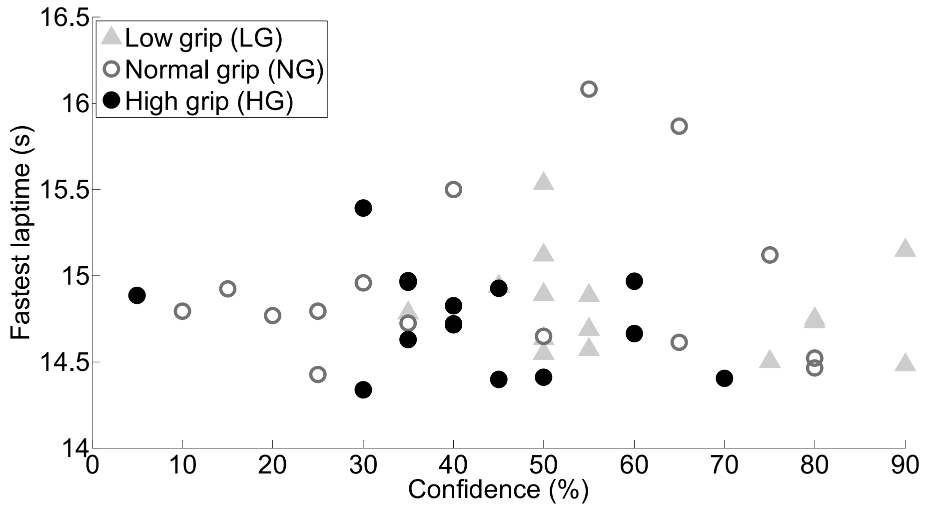

Figure 6. Fastest lap time during Retention 2 session (driven with the Formula 1 car) as a function of the response to Statement 3 of the Confidence questionnaire ("I think I had a faster lap time than the average participant in my group") for all 45 participants.

7.4 Discussion

This simulator-based study investigated the effect of the tire-road friction coefficient on the training effectiveness of a car racing task. Three groups of inexperienced racecar drivers practiced with different grip and were tested on their ability to drive the fastest possible lap with normal tire-road friction coefficient. Practicing with low grip level resulted in slower lap times, higher self-confidence, and less frustration. The higher self-confidence of the low-grip group was as expected, but apparently this higher self-confidence did not lead to faster lap times. Instead, these participants thought they were faster than the average participant in their group, were not as frustrated, and also crashed less in the opening laps. Thus, it appears that the low-grip group was more complacent and drove at a more comfortable speed, further from the grip limit of the normal-grip car.

The effects of the low-grip training were strongest in the opening laps of the first retention session (see Figure 3), where the lap times were very slow and the low-grip group had less rejected laps. For racecar driver training, it would be relevant to study the longer-term retention effects as well. Considering that sleep has an important role in consolidation of memory, we expect a skill improvement for all groups on the following day (cf. Walker, Brakefield, Morgan, Hobson, & Stickgold, 2002). Previous research in a driving simulators showed that training effects are retained overnight, although the effect sizes between groups were attenuated as compared to immediate retention (De Groot, De Winter, López-García, Mulder, & Wieringa, 2011).

Previous research showed that there is a significant correlation between the lap times driven in the racing simulator and lap times driven in real-world practice sessions (De Groot, De Winter, Mulder, & Wieringa, 2011a). However, care must be taken to extrapolate the results of this simulator-based study to the real world. The risk perception of drivers in the simulator was low (overall average 29% to Statement 1 of the confidence questionnaire), which is without doubt much lower than in reality. One of the advantages of the driving simulator, driver safety, might also cause behavior that is not realistic: many laps were rejected because the car hit the wall or left the tarmac. On the other hand, these errors are part of a new method of training; a method only possible in the safe simulator environment.

Part 3

Racing simulator validity and controllability

Chapter 8. Car racing in a simulator: validation and assessment of brake pedal stiffness

Abstract

Car racing is a mentally and physically demanding sport. The track time available to train drivers and test car setups is limited. Race car simulators offer the possibility of safe, efficient, and standardized human-in-the-loop training and testing. We conducted a validation study of a racecar simulator by correlating the fastest lap times of 13 drivers during training events in the simulator with their fastest lap times during real-world race events. The results showed that the overall correlation was .57 (n = 13). Next, the effect of brake pedal stiffness (soft: 5.8 N/mm versus hard: 53.0 N/mm) on racing performance was investigated in the simulator. Brake pedal stiffness may have an important effect on drivers' lap times, but it is impractical to manipulate this variable on a racecar during a real-world test session. Two independent experiments were conducted using different cars and tracks. In each experiment participants (n = 6 in Experiment 1 and n = 9 in Experiment 2) drove alternately with the soft and hard pedal in eight 20-min sessions (Experiment 1) or six 15-min sessions (Experiment 2). Two hypotheses were tested: 1) the soft pedal yields faster cornering times for corners that include a long brake zone, and 2) the hard pedal yields more high-frequency brake forces. Experiments 1 and 2 confirmed the second hypothesis but not the first. Drivers were highly adaptable to brake pedal stiffness, and the stiff pedal elicited higher and more high-frequent brake pedal forces. It is concluded that the racing simulator is a valuable tool for testing and driver assessment.

De Groot, S., De Winter, J. C. F., Mulder, M., & Wieringa, P. A. (2011a). Car racing in a simulator: Validation and assessment of brake pedal stiffness. *Presence: Teleoperators and Virtual Environments, 20*, 47–61.

8.1 Introduction

Car racing is a physically and mentally demanding sport. It not only requires physical strength and rapid cognitive processing (Jacobs, Olvey, Johnson, & Cohn, 2002; Klarica, 2001), it is also a high-risk sport (Minoyama & Tsuchida, 2004). A racing car requires a team of personnel for maintenance and preparation prior to each track outing, and due to constraints posed by regulations, time, and financial budgets, racing teams often have limited possibilities to train and evaluate drivers as well as car setups. Furthermore, varying track, tire, and vehicle properties during test sessions restrict the possibilities to perform valid comparisons and to draw appropriate conclusions.

Not surprisingly, racing teams often rely on computer-based tools to develop their cars, to think out racing strategies, and to test car setups. For example, Casanova, Sharp, and Symonds (2002) used a simulation to determine the optimal longitudinal location of the center of mass of a Formula One racing car. This simulation used an automatic controller rather than an actual human driver, and Casanova et al. acknowledged that the optimal location of the mass center might be different for human drivers. It’s only possible to evaluate human-machine interaction and effects of driver adaptation through human-in-the-loop simulation.

Olsen, Page, and Vulovic (2008) described the development of a human-in-the-loop racing simulator to be used as an engineering tool and as a driver-training tool for students competing in Formula Student design competitions. For rally driver training, Walker, Dawson, and Ackland (2001) used a simulator to increase drivers’ skill and physical endurance. Cipelli, Schiehlen, and Cheli (2008) describe the development of a simple car model which can be used for human-in-the-loop research. More generally, in the field of driver testing for rehabilitation purposes, it is increasingly recognized that driving simulators offer a promising alternative to on-the-road methods of driver assessment (Kraft, Amick, Barth, French, & Lew, 2010). Furthermore, in many sports other than racing, athletes and coaches increasingly rely on human-in-the-loop simulation (Liebermann et al., 2002).

Indeed, the potential advantages of a (racecar) simulator are many. Simulators offer the possibility of standardized and relatively inexpensive training and testing. A simulator is inherently safe, and time-varying effects such as fuel use and tire wear can be disabled, providing better experimental control compared to training and testing on a real track. Furthermore, a simulator allows for highly accurate measurements of the vehicle state, something that requires expensive sensors in real racing cars. In addition, testing conditions and car setups can be adapted quickly, and testing sessions will not be delayed because of crashes and car damage.

Clearly, the benefits of simulator-based training and assessment only hold when the drivers’ behavior in the simulator has predictive value and when the drivers’ skills gained in the simulator transfer to the real race track. Therefore, validation should be a crucial component of any sports simulator. One way to assess the validity of a simulator is to statistically associate the performance measures obtained in the simulator with those obtained in reality (e.g., Shechtman, Classen, Awadzi, & Mann, 2009; De Winter et al., 2009).

In this study, we first provided a validation study of our racecar simulator by correlating the lap times measured in the simulator with the lap times driven on the real race track by the same drivers. Second, we aimed to exploit the aforementioned advantages of a racecar simulator by testing the effect of brake pedal stiffness on the driver’s the lap times. It is impractical to manipulate brake pedal stiffness during a test session on a real track due to variable track, tire, and car conditions, long required times to manipulate the brake system, and safety concerns. Anecdotically, considerable variation in pedal stiffness exists between cars and between drivers' preferences. The Delft University of Technology (DUT) Formula Student Racing Team has built cars with brake pedal stiffness varying a magnitude of ten over the last three years (see also Figure 1), and it was still unknown which pedal stiffness results in the best braking performance.

The braking system of racing cars can be characterized by the relationship between three variables: brake pedal displacement, brake pedal force, and vehicle deceleration. The relationship between pedal force and pedal displacement determines the stiffness of the brake pedal. The relationship between pedal force or displacement and vehicle deceleration is a measure of the sensitivity of the brake pedal. Note that electronic assistance systems such as anti-lock braking systems (ABS) or electronic brake-force distribution (EBD) are generally not permitted in car racing. To brake as effectively as possible, the driver should use the grip offered by the car and tires to its maximum potential. This can be achieved by late braking and adopting a threshold braking technique; in other words, rapidly reaching the maximum possible deceleration (buildup phase) and then modulating the brake pressure to keep the slip ratio of the tires at the optimal value (modulating phase).

When pedal displacements are larger (i.e., with a softer pedal), more proprioceptive feedback is provided to the driver about his foot position. The theory of sensory integration predicts that a person will integrate all available sensory information to minimize a target error (Ernst & Bülthoff, 2004; Mugge, Schuurmans, Schouten, & Van der Helm, 2009). Our first hypothesis was that a softer brake pedal yields faster cornering times if the corner includes a long brake zone (implying a relatively long modulating phase). The second hypothesis was that a stiffer brake pedal would cause more high-frequency brake force inputs than a softer pedal, because the driver's foot does not have to be moved as much during braking. The two hypotheses were tested in two independent experiments in the driving simulator. Experiment 1 was a relatively simple task, with a Formula-Student-type trainer vehicle, an artificial racetrack, relatively slow driving speeds (maximum speed 90 km/h), and forgiving tires. Experiment 2 was a more complex task conducted by semi-professional drivers, relatively high driving speeds (maximum speed 230 km/h) on a realistic circuit, and less forgiving tire characteristics. The reason for conducting the second experiment was to investigate the generalizability of the results of the first experiment in a more complex context.

8.1.1 Simulator hardware and software

The simulator was constructed by the first author from a Formula Renault 2.0 chassis. The simulator was fixed-base and therefore provided no vestibular feedback. To control the virtual car, drivers had to steer, brake, and apply throttle. During the validation study and during Experiment 2, drivers had to manually change gears as well. The original steering rack, throttle system (including cable and spring mechanism), brake system (including master cylinders and brake calipers with hydraulic lines), and sequential gear lever (including cable and mechanical force feedback) were the control devices of the simulator. These original control elements were used in order to get the control feel as close to reality as possible. A force feedback motor (Thrustmaster RGT off-the-shelf product) was connected to the original steering shaft through a gear-belt drive.

We used rFactor version 1.255 as simulation software. rFactor is generally recognized as one of the most accurate racing simulators available (Wikipedia, 2010), providing advanced models of vehicle dynamics, tire characteristics, and steering wheel force feedback. Furthermore, rFactor provides ample possibilities of customization and modification for scientific applications (Weinberg & Harsham, 2009; Weinberg, Harsham, Forlines, & Medenica, 2010). To ensure constant driving conditions, tire-wear and fuel consumption were disabled in the validation study and in both experiments.

The simulated environment was displayed on a 19-inch monitor positioned in front of the participant at a distance of 90 cm from the driver's face. The screen resolution was 1,280 x 1,024 pixels, resulting in a vertical field of view of 18 degrees and a horizontal field of view of 24 degrees. One may argue that the 19-inch monitor is too small to obtain a high degree of immersion, considering the previous research on the relationship between field of view and presence (e.g., Hendrix & Barfield, 1996; Lin, Duh, Parker, Abi-Rached, & Furness, 2002). However, in our brief post-experiment interviews, participants reported no problem at all with the visual display or with the degree of immersion. This is in line with research of Särkelä, Takatalo, Komulainen, Nyman, and Häkkinen (2004) which reported a relatively high degree of presence amongst participants racing with a 21-inch monitor positioned at 1 m in front of their face. Indeed, racing is a cognitively demanding activity, and previous research has indicated that mental workload is positively correlated with the degree of presence (Ma & Kaber, 2006). Racing drivers are usually highly immersed in their task, as demonstrated by a repeatable pattern of looking to the track edges without consideration of external stimuli (Land & Tatler, 2001).

The visual frame rate of the simulation was 60 Hz. The steering wheel was not visualized in the virtual environment, contrary to the cockpit and front tires. Sound was provided by means of four speakers positioned around the participant. No other vehicles were present on track during any driving session. The data logging took place at a frequency of 100 Hz.

8.1.2 Validation study: simulator-reality correlations

The validity of the simulator was investigated by comparing the fastest lap times obtained during training events in the simulator with the fastest lap times on the same race tracks during real-world race events. During the real-world race events, the overall fastest lap time of the free practice and qualifying sessions was taken into consideration. Note that the race itself was not included because during races, the drivers' task is to achieve the best possible position at the finish line, and not necessarily to drive the fastest lap time per se. In the simulator and in the real world, feedback on performance after each driving session was based on the drivers' fastest lap data, consisting of speed, engine rpm, and control input signals, compared to the data of a competitive reference lap.

The first simulator event was conducted with seven Formula Renault drivers from one real-world Formula Renault racing team. Each driver trained with a virtual Formula 3 car on the tracks of Zandvoort (Netherlands) and Hockenheim (Germany) on the same day. The drivers completed three 15-min simulator sessions per track. The mean total number of simulator laps per driver was 24.7 on Zandvoort and 25.0 on Hockenheim. During the corresponding real-world event a mean total of 28.7 laps per driver were completed on Zandvoort and 34.4 laps on Hockenheim.

The second simulator event was performed with five Formula Renault drivers, four of which also participated in the first training event. They trained with a virtual

Formula 3 car on the track of Most (Czech Republic), which was unknown to all of them. The data obtained during this training event were also part of Experiment 2 of this article. The drivers completed six 15-min simulator sessions. The mean total number of laps per driver was 44.8 in the simulator and 81.8 during the corresponding real-world event. Such a large number of real-world laps were possible because the drivers had a full day of testing time before the start of the actual race event.

The third simulator event was performed with five drivers from one real-world Renault Mégane racing team competing in the European Renault Mégane championship. They trained with a virtual Renault Mégane racecar on the Motorland Aragón track (Spain). The drivers completed three 15-min simulator sessions. The mean total number of laps per driver was 19.2 laps per driver in the simulator and 44.4 during the corresponding real-world event.

Table 1 shows the results of the validation study. It can be seen that the variability between drivers was smaller in reality than in the simulator. Because we used a Formula 3 car to train Formula Renault drivers, the lap times in the simulator were faster than in reality during the Zandvoort, Hockenheim, and Most sessions. A Formula 3 car is normally faster than a Formula Renault car. We used a Formula 3 car because a Formula 3 model was supplied with the rFactor software (rF3), while a Formula Renault model was not readily available. Formula 3 and Formula Renault cars are both formula cars with a considerable amount of downforce, and similar mass, power, and tire grip. In contrast, for the Mégane drivers the lap times in reality were faster than in the simulator. The inter-track correlations of the simulator tracks were comparable to the inter-track correlation of the real tracks (all correlations > .75), indicating that both the simulator and the real track provided reliable measurement data. The correlations between the fastest lap times in the simulator and the fastest lap times on the real track were .63 (p = .13, n = 7), .60 (p = .16, n = 7), .20 (p = .75, n = 5), and .63 (p = .25, n = 5) for Zandvoort, Hockenheim, Most, and

Table 1. Mean fastest lap time, standard deviation of fastest lap time, and pairwise correlation matrix between simulator and real-world fastest lap times.

				Simulator				Real		
		M lap time (s)	*SD* lap time (s)	Zandvoort	Hockenheim	Most	Motorland	Zandvoort	Hockenheim	Most
Simulator	Zandvoort	95.95	2.98							
	Hockenheim	95.46	3.00	.99 (7)						
	Most	92.83	1.35	.80 (4)	.82 (4)					
	Motorland	129.38	3.86	-	-	-				
Real	Zandvoort	97.59	0.78	.63 (7)	.61 (7)	.01 (4)	-			
	Hockenheim	99.09	0.59	.62 (7)	.60 (7)	.38 (4)	-	.75 (7)		
	Most	93.11	0.50	.78 (4)	.76 (4)	.20 (5)	-	.89 (4)	.92 (4)	
	Motorland	126.03	1.91	-	-	-	.63 (5)	-	-	-

A value of 1 corresponds to a perfect linear relationship between the driver's lap times in the simulator and reality; a value of 0 means that there is no linear relationship. The number of pairs of participants used to calculate the correlation coefficients is listed between parentheses. Thirteen drivers participated in this validation study. Seven drivers participated on Zandvoort and Hockenheim, five drivers participated on Most (four of these also participated on Zandvoort and Hockenheim), and a further five drivers participated on the Motorland track.

Motorland, respectively.

Next, we standardized the fastest lap times per race track by subtracting the mean and dividing by the standard deviation, such that it became possible to aggregate the lap times across different tracks. We then calculated the drivers' average standard scores of the simulator tracks and the drivers' average standard scores of the real race track. The correlation between simulator and reality was .57 (p = .044, n = 13).

8.2 General method of experiments 1 and 2

8.2.1 Independent variable

A low and a high brake pedal stiffness were investigated, referred to as Soft and Hard, respectively. Soft had a stiffness of 5.8 N/mm, Hard a stiffness of 53.0 N/mm. These values were based on measurements conducted on the Delft Formula Student cars DUT06 and DUT08 (see Figures 1 and 2). The measurements were conducted with a linear potentiometer and a load cell temporarily mounted onto the brake pedal, both in the simulator and in the two Formula Student cars. Note that the load cell and potentiometer were used only for measurement of the brake pedal characteristics as shown Figures 1 and 2; they were not used in any of the driving sessions.

During driving sessions, the force exerted on the pedal was continuously measured using strain gauges attached to the brake pedal itself, and this force was used as an input to the rFactor software. The maximum input as supplied to the software (i.e., 100% brake input) corresponded to a brake pedal force of 375 N. Between each driving session the brake pedal was modified from Soft to Hard and vice versa. The difference in stiffness was obtained by replacing the original brake pad with a softer material.

8.2.3 Tire grip and slip characteristics

The vehicle model which was used in Experiment 1 differed from the model used in Experiment 2. Figure 3 shows the brake input of the vehicle model (corresponding linearly to the measured brake pedal force) versus the longitudinal deceleration of both simulated vehicles, indicating that the deceleration/pedal-force gain was larger

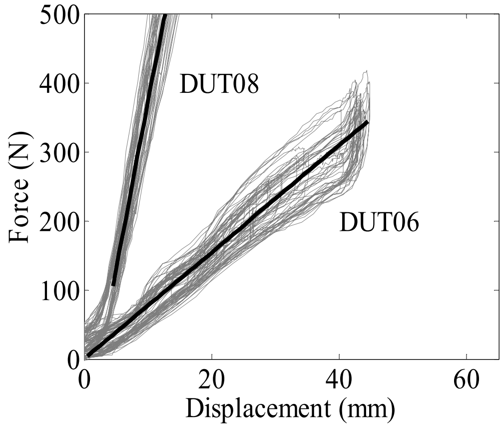

Figure 1. Delft University of Technology Formula Student car brake pedal characteristics. DUT06 and DUT08 are the 2006 and 2008 car models, respectively. Grey lines show pedal force versus displacement.

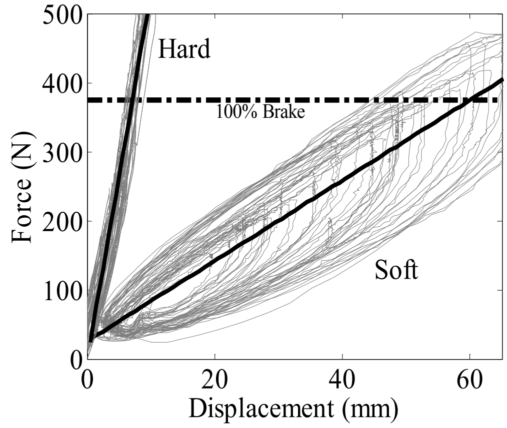

Figure 2. Driving simulator brake pedal characteristics. Grey lines show pedal force versus displacement measurements, black lines represent a linear fit. The black dash/dotted line represents maximum (100%) brake input provided to the simulator.

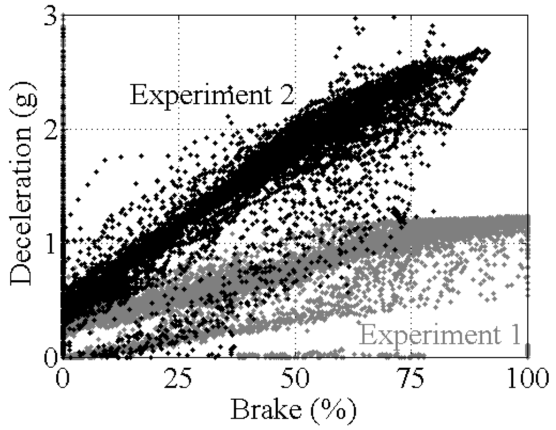

Figure 3. Deceleration versus brake depression for the rTrainer (Experiment 1) and rF3 (Experiment 2).

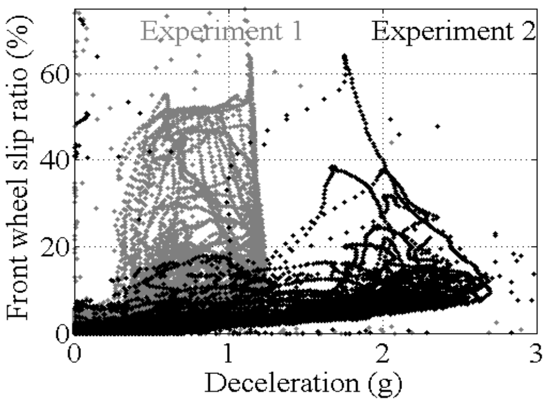

Figure 4. Front wheel slip ratio versus deceleration for the rTrainer (Experiment 1) and rF3 (Experiment 2).

in Experiment 2. Figure 4 shows the slip ratio versus vehicle deceleration. The slip ratio was defined as the ratio between the mean front wheel speed and the vehicle ground speed (Milliken & Milliken, 1995). The maximum deceleration of the vehicle used in Experiment 1 remained almost constant for increasing front wheel slip ratio, whereas the vehicle used in Experiment 2 generated maximum grip at around a slip ratio of 10% and lost deceleration capability with increasing wheel slip ratio.

8.3 Experiment 1: Brake pedal stiffness in a simple racing task

8.3.1 Participants and instructions

Six men participated in Experiment 1. Five participants had previously driven Formula Student cars on Formula Student or Formula SAE competitions. The sixth participant had completed a racing course on a real-world track. Participants were all TU Delft students or employees, and their mean age was 26.7 years (*SD* = 4.1). The first two authors of this paper were amongst the six participants and were not the experimenters.

Participants were instructed to keep at least two wheels on the asphalt, and their goal was to drive the fastest possible lap within each of the eight 20-min driving sessions. After the completion of each lap, a synthesized voice reported the participant's lap time. No other feedback than the driver's own lap times was provided during the experiment. After each session, participants were asked to step out of the simulator and take a break of approximately 10 min (during which the brake configuration was altered by the experimenters).

8.3.2 Race track

A race track was developed which included low-to-medium speed turns, typical for Formula Student competitions. The developed track consisted of five turns and one chicane. The track was 7 m wide and 1,160 m long. Section 1 included a 180 degree hairpin, Section 2 a 90 degree right-hander, Section 3 a 90 degree left turn, Section 4 a 120 degree left turn, and Section 5 a fast 60 degree left turn. The chicane was located between Sections 4 and 5. Before every turn, four 10-meter-spaced signs were placed alongside the track to help drivers judge their braking location. To illustrate the track characteristics, Figure 5 shows the track sections, vehicle ground

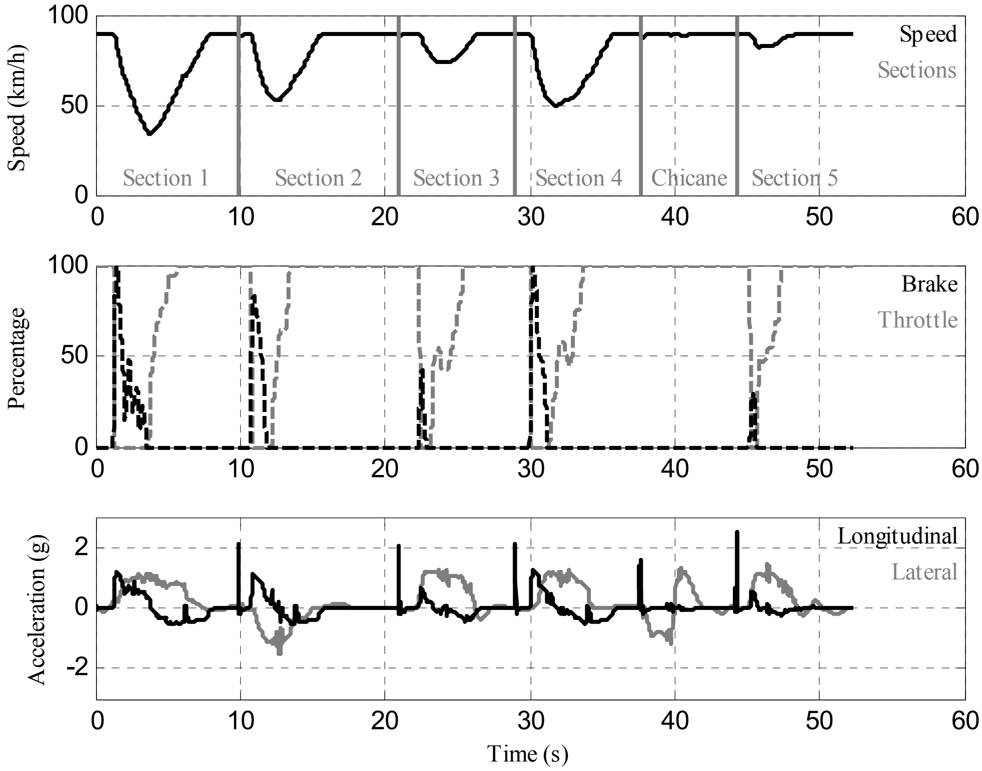

Figure 5. Speed, sections, throttle, brake, and vehicle accelerations of the fastest recorded lap (52.26 s; Experiment 1). The positive spikes in the longitudinal acceleration signal were caused by purposefully placed bumps, and used for automatic calculation of the section times.

speed, brake and throttle percentages, and lateral and longitudinal acceleration of the fastest recorded lap. It can be seen that Sections 1, 2, and 4 involved hard braking, whereas Sections 3 and 5 required only a "tap" on the brake pedal. The chicane is not visible in the longitudinal acceleration signal, as it was possible to take this combination of turns without releasing the throttle.

8.3.3 The vehicle model

The vehicle model used was the so-called rTrainer vehicle as available in the rFactor software. This rear-wheel drive formula racing car had a mass of 630 kg and a maximum engine power of 85 kW. The car generated little down-force and the maximum lateral and longitudinal acceleration was about 1.2 g. The gearbox and clutch were automated during the experiment, and there was no ABS. The maximum speed of the vehicle was set at 90 km/h, which was normally reached on each straight. In this way the exit speed of a section could not influence the approach speed of the next section, therefore guaranteeing statistical independence of each section time.

8.3.4 Results

Section times. Table 2 shows the fastest section times per participant. Recall that our first hypothesis was that Soft would be more beneficial in sections containing a long brake zone (i.e., Sections 1, 2, and 4 in Table 2). Table 2 reveals that four participants were indeed fastest with Soft in Section 1. However, only one participant was fastest with Soft in Section 2. In Section 4, Soft and Hard were equally frequent

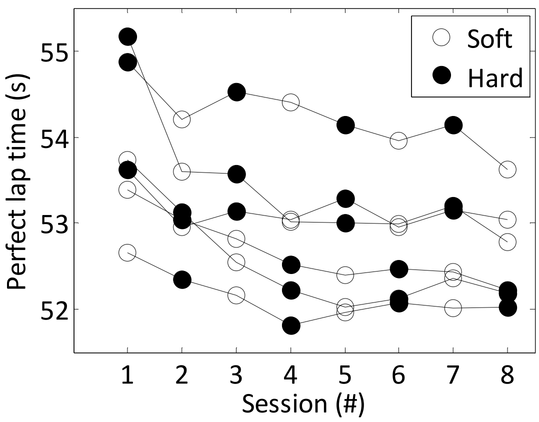

Figure 6. Perfect lap time as a function of session for all participants (Experiment 1). Each line corresponds to a participant. The perfect lap time is defined as the sum of the participant's fastest section times within one driving session.

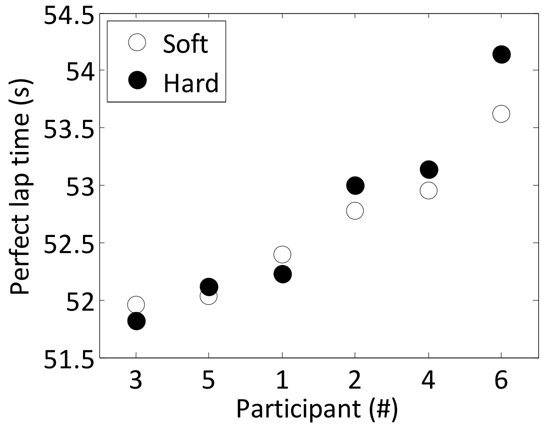

Figure 7. Fastest perfect lap time of all sessions with Soft and Hard per participant (Experiment 1).

between the participants. None of these effects was statistically significant, and so the first hypothesis was not confirmed.

Perfect lap times. Figure 6 shows the perfect lap times per participant as a function of session number. Participants considerably reduced their perfect lap time with increasing session number. Comparing the mean of the perfect lap times of Sessions 1 and 2 (to exclude possible pedal condition effects) with the mean of the participants' perfect lap times of Sessions 7 and 8 revealed a significant learning effect (*M* Sessions 1 & 2 = 53.56 s, *M* Sessions 7 & 8 = 52.77 s, t = 4.41, p = .007). Herein, t is the t statistic, a statistic whose sampling distribution is a Student's t

Table 2. Fastest section times as a function of participant, brake pedal, and section number (Experiment 1).

Participant	Pedal	Section 1	2	3	4	5
3	Soft	9.70	10.98	8.08	8.52	8.04
	Hard	**9.69**	**10.89**	**8.04**	**8.51**	**8.01**
5	Soft	9.77	10.95	**8.07**	8.56	8.10
	Hard	9.77	**10.89**	8.09	**8.54**	**8.08**
1	Soft	**9.83**	11.13	**8.07**	**8.59**	8.11
	Hard	9.87	**11.05**	8.08	8.61	**8.03**
2	Soft	**9.99**	11.17	8.20	**8.66**	**8.15**
	Hard	10.05	**11.16**	**8.17**	8.75	8.22
4	Soft	**9.90**	**11.21**	**8.18**	**8.63**	8.31
	Hard	9.95	11.31	8.19	8.69	**8.27**
6	Soft	**10.10**	11.37	**8.38**	8.79	**8.41**
	Hard	10.19	**11.33**	8.39	**8.77**	8.66

The chicane was not listed because this section could be taken without braking. A section time of 6.57 s was substituted for all participants.

distribution, and p represents the probability of observing the given means by chance if the means would actually be equal; p was calculated using a paired t test.

Figure 7 shows the fastest perfect lap times per participant for Soft and Hard. It can be seen that the within-subjects variability as a result of the pedal configuration is much smaller than the between-subjects variability. The group means did not differ significantly (M Soft = 52.62 s, M Hard = 52.74 s, $t = -1.13$, $p = .308$).

Brake input. All participants had their brake fully depressed for a larger percentage of the time with Hard (M Soft = 0.48%, M Hard = 0.98%, $t = -4.93$, $p = .004$). This indicates that Hard promoted braking with higher forces (note that the force necessary to reach 100% brake input was equal for Soft and Hard). All drivers had more high-speed-increasing brake input with Hard, supporting the second hypothesis (M Soft = 0.19%, M Hard = 0.33%, $t = -2.45$, $p = .058$). They had more high-speed-releasing brake input with Hard as well (M Soft = 0.13%, M Hard = 0.21%, $t = -2.54$, $p = .052$), also supporting the second hypothesis.

8.4 Experiment 2: Brake pedal stiffness in a complex racing task

8.4.1 Participants and instructions

Nine men participated in Experiment 2. Six participants were Formula Renault drivers (five of which also participated in the validation study above), one was a DUT Formula Student driver which did not participate in Experiment 1, and the remaining two were the first two authors of this article. All participants had prior experience with the racing simulator, and their mean age was 20.4 years (SD = 5.5).

Participants were instructed to keep at least two wheels on the asphalt and their goal was to drive the fastest possible lap within each of the six 15-min driving sessions. After the completion of each lap, a synthesized voice reported the participant's lap time. After each session, participants were asked to step out of the simulator and take a break of approximately 10 min, during which the brake configuration was altered by the experimenters. In contrast to Experiment 1, drivers were provided with a comparison of their last session's fastest lap speed and control input signals against a competitive reference lap.

8.4.2 Race track

Experiment 2 was conducted on a simulated version of the Most track (Czech Republic), which was unknown to all participants. The track was between 12 and 15 m wide and 4,170 m long. Section 1 comprised a slow chicane after the start/finish straight, Section 2 consisted of a combination of three medium-speed turns (with two braking actions) and Section 3 was a fast right turn. Section 4 consisted of another medium-speed turning combination. Section 5 consisted of the final two 90 degree right turns before the start/finish straight. Figure 8 shows the track sections, the vehicle ground speed, brake and throttle inputs (on a scale from 0 and 100%), and lateral and longitudinal acceleration of the fastest recorded lap. Sections 1, 2 and 5 involved severe braking, whereas Sections 3 and 4 required less severe braking.

8.4.3 The vehicle model

The vehicle model used was the rF3 vehicle as available in the rFactor software. This rear-wheel drive formula racing car has a mass of 550 kg and a maximum engine

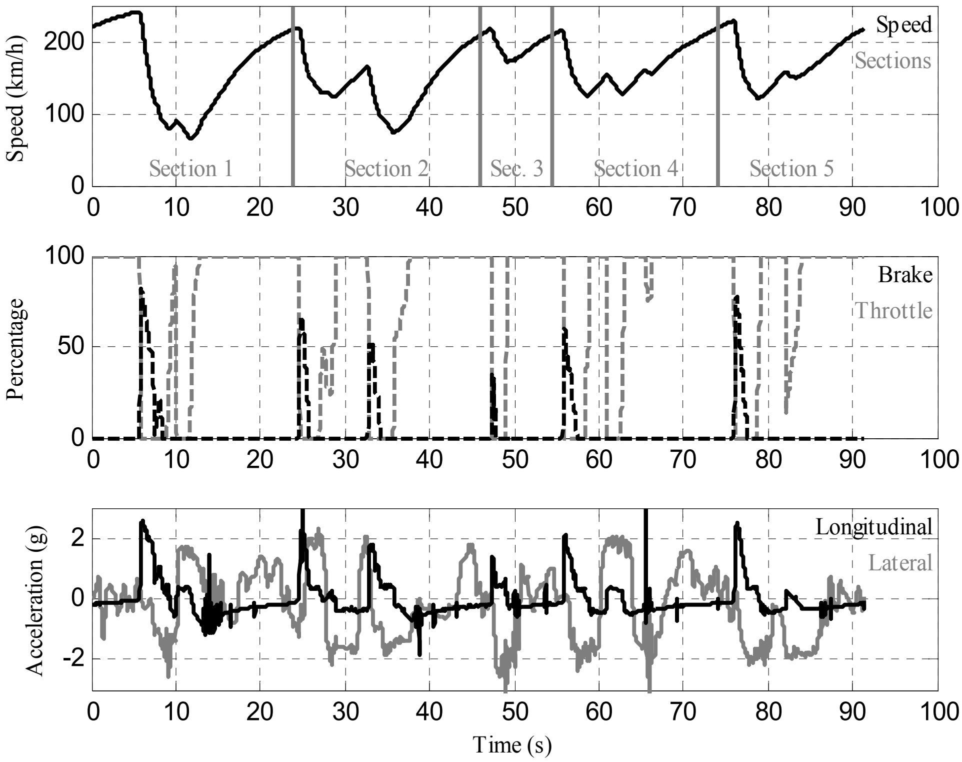

Figure 8. Speed, sections, throttle, brake, and accelerations of the fastest recorded lap (91.42 s; Experiment 2).

power of 162 kW. The rF3 generated down-force and the maximum lateral acceleration was about 2.4 g in Section 3. The maximum longitudinal acceleration, during braking in Section 1, was 2.5 g. At lower speed the accelerations were limited to about 1.7 g. The sequential gearbox was manually operated and the clutch was automated.

8.4.4 Results

Section times. Table 3 shows the fastest section times for each participant. According to the first hypothesis, Soft would be more beneficial in Sections 1, 2 and 5, which contained a long brake zone. It can be seen that three participants were faster with Soft in Section 1, five in Section 2 and six participants were faster with Soft in Section 5. As in Experiment 1, no significant section time effects between pedals were found.

Perfect lap times. Participants reduced their perfect lap time with increasing session number (see Figure 9). Comparing the mean of the perfect lap times of Sessions 1 and 2 to the mean of the perfect lap times of Sessions 5 and 6 revealed a significant difference (*M* Sessions 1 & 2 = 93.68 s, *M* Sessions 5 & 6 = 91.97 s, t = 4.62, p = .002). Figure 10 shows the fastest lap time of all participants and for both pedal configurations. As in Experiment 1 there was no significant difference between the means of the participants' fastest perfect lap time for the different pedals (*M* Soft = 91.77 s, *M* Hard = 91.89 s, t = −1.12, p = .295).

Brake input. As in Experiment 1, all participants had their brake fully depressed for a larger percentage of the time with Hard (*M* Soft = 0.71%, *M* Hard = 1.08%, t = −2.80, p = .023). All drivers had more high-speed brake input using Hard,

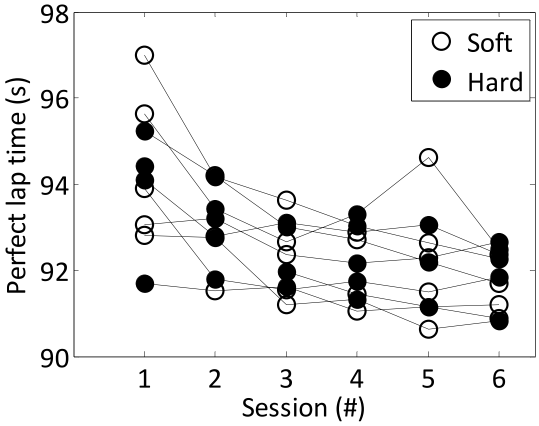

Figure 9. Perfect lap time as a function of session for all participants (Experiment 2). Each line corresponds to a participant. The perfect lap time is defined as the sum of the participant's fastest section times within one driving session.

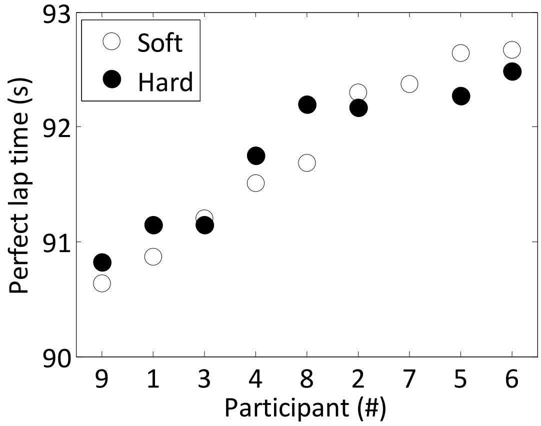

Figure 10. Fastest perfect lap time of all sessions with Soft and Hard per participant (Experiment 2).

confirming the second hypothesis (*M* Soft = 0.29%, *M* Hard = 0.35%, $t = -2.68$, $p = .028$). The brake derivative for the brake release revealed no significant difference between pedals (*M* Soft = 0.16%, *M* Hard = 0.18%, $t = -1.19$, $p = .270$).

Table 3. Fastest section times as a function of participant, brake pedal, and section number (Experiment 2).

		Section				
Participant	Pedal	1	2	3	4	5
9	Soft	**23.68**	**22.30**	**7.86**	**17.31**	19.42
	Hard	23.81	22.43	7.90	17.33	**19.32**
1	Soft	23.84	**22.38**	7.93	**17.22**	**19.37**
	Hard	**23.73**	22.53	**7.92**	17.41	19.50
3	Soft	24.02	**22.39**	**7.82**	**17.35**	19.57
	Hard	24.02	22.44	7.87	17.44	**19.37**
4	Soft	24.08	22.44	7.91	**17.30**	**19.50**
	Hard	24.08	**22.37**	**7.90**	17.43	19.58
8	Soft	**23.99**	**22.58**	**7.93**	**17.48**	**19.59**
	Hard	24.15	22.90	7.94	17.50	19.72
2	Soft	**23.81**	22.90	8.01	17.66	**19.47**
	Hard	24.05	**22.87**	**7.90**	**17.58**	19.72
7	Soft	24.25	**22.41**	**8.17**	**17.64**	**19.49**
	Hard	**24.09**	22.97	8.23	17.82	19.67
5	Soft	24.10	22.94	8.06	17.87	19.66
	Hard	**24.06**	**22.84**	**8.05**	**17.63**	**19.53**
6	Soft	24.21	23.09	7.96	**17.69**	**19.72**
	Hard	**23.86**	**22.85**	7.96	17.77	19.75

8.5 Discussion

In this article, we provided a validation study of a racecar simulator showing that drivers' fastest lap times in the simulator correlated well with the fastest lap times during race events on the real-world tracks, a significant effect. This supports the predictive value of the measures obtained in the driving simulator. Next, we investigated the effects of brake pedal stiffness on driver performance in the racecar simulator. In two experiments of different task complexity, we showed that the section times were not significantly affected by brake pedal stiffness, indicating that drivers adapted well to the variations in the human-machine interface. We did find that the stiffer pedal promoted faster brake inputs and higher pedal forces.

In addition to the effects of brake pedal stiffness, two stronger effects were observed. First, a learning effect was found in the simulator and on the real track (data not shown). Such a learning curve is not unique to car racing. For other tasks, such as reaction times (one of the simplest possible tasks) and rolling cigars (extreme case of ongoing improvement after thousands of trials), practice effects occur as well (Crossman, 1959; Jensen, 2006, p. 46). These learning effects may have complicated the validation study in which we correlated the drivers' best lap times between the simulator and reality. However, considering that all drivers drove approximately an equal number of laps and that only the drivers' fastest lap time was taken into consideration, it can be assumed that the learning effect did not notably interact with the observed correlations. Second, large individual differences were found. The learning effect and the large individual differences imply that it is important to select and train race drivers in order to obtain fast lap times.

Participants showed a marked degree of adaptability to the brake pedal stiffness. In both experiments, the pedal displacements at 100% brake input differed by almost a factor 10, but the lap time differences were not statistically significant. Previous manual control experiments have also shown that humans are able to effectively adapt their behavior across a wide range of control-display gains (Casiez, Vogel, Balakrishnan, & Cockburn, 2008; Huysmans, De Looze, Hoozemans, Van der Beek, & Van Dieën, 2006; McRuer & Jex, 1967; Van Doorn, Unema, & Hendriks, 2005). Our results resemble those of Segel and Mortimer (1970) who systematically varied the brake pedal stiffness of a passenger car with an unassisted braking system, while keeping the deceleration versus pedal-displacement function constant (i.e., a position-control pedal). Their results also showed that drivers were highly adaptable to varying pedal configurations. The minimum stopping distance depended predominantly on the individual driver capabilities, initial speed, and surface friction coefficient, and not as much on the pedal stiffness, even though the latter was varied by over a factor 16. Contrary to the work of Segel and Mortimer (1970), however, this study investigated the effect of pedal stiffness while keeping the deceleration/pedal-force gain constant. In other words, the force needed for the driver to decelerate the car remained identical while the pedal displacements varied (i.e., a force-control pedal). This may provide a more fair comparison between varying levels of stiffness, because the physical effort of the driver remains constant. Segel and Mortimer (1970) found an optimum deceleration/pedal-force gain for a given grip level of the tire-road interface, and it would be interesting to investigate whether other gains (i.e., by increasing or decreasing the brake pedal force necessary to decelerate the vehicle) cause a performance improvement or decrement. A higher deceleration/pedal-force gain would result in less required effort, but also less proprioceptive feedback, whereas oppositely, a lower deceleration/pedal-force gain would result in more physical effort and more proprioceptive feedback.

Finally, the simulator fidelity needs some discussion. Although the characteristics of the pedals as well as the vehicle models were realistic and representative, there may have been unrealistic aspects to the driving simulation as well. In the real world, decelerations induce an extra pedal force as a result of the mass of the driver’s leg. Furthermore, the driver is provided with tactile and vestibular feedback, providing direct information about the car’s deceleration. In our experiment, the drivers inferred the deceleration from visual and auditory cues only.

Our results indicated that the variability of lap times was larger in the simulator than in reality. In ongoing work, we increased the simulator fidelity by installing a larger visual display and by including improved force feedback on the steering wheel (see Figure 11 vs. Figure 12). We expect that lap time variability within and between drivers will decrease as a result. This hypothesis will be tested in future experiments.

In conclusion, this study showed that lap times during simulator practice sessions were predictive of lap times during real-world car racing events. Drivers are highly adaptable to varying brake pedal stiffness, but a stiffer pedal elicits more rapid brake force variations and higher brake pedal forces. The racing simulator was found a viable alternative for on-track testing, offering safe driving conditions, valid measurements, and efficient experimental control.

Figure 11. Racing simulator as used in the validation study as well as Experiments 1 and 2. The brake calipers were masked by plastic bags so that the participant could not see the brake system during Experiments 1 and 2.

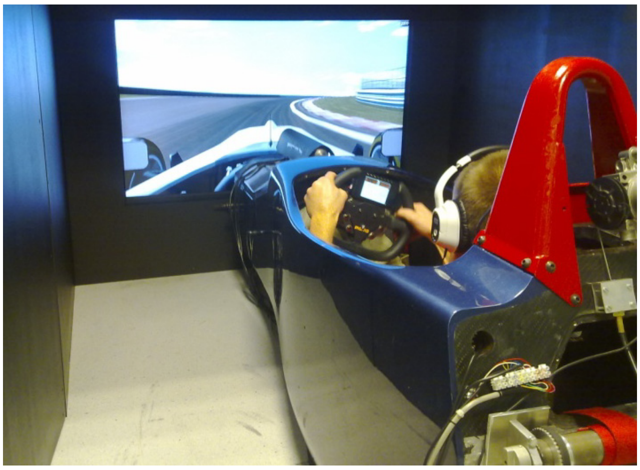

Figure 12. Racing simulator at the time of publication. In comparison to the experimental version (Figure 11), a large (52-inch) screen and side covers were installed to improve presence, as well as an improved steering wheel force feedback system, headphones, and a seatbelt tensioning system.

Chapter 9. The effects of control-display gain on performance of race car drivers in an isometric braking task

Abstract

To minimise lap times during car racing, it is important to build up brake forces rapidly and maintain precise control. We examined the effect of the amplification factor (gain) between brake pedal force and a visually represented output value on a driver's ability to track a target value. The test setup was a formula racing car cockpit fitted with an isometric brake pedal. Thirteen racing drivers performed tracking tasks with four control-display gains and two target functions: a step function (35 trials per gain) and a multisine function (15 trials per gain). The control-display gain had only minor effects on root mean-squared error between output value and target value, but it had large effects on build-up speed, overshoot, within-participants variability, and self-reported physical load. The results confirm the hypothesis that choosing an optimum gain involves balancing stability against physical effort.

De Winter, J. C. F., & De Groot, S. (2012). The effects of control-display gain on performance of race car drivers in an isometric braking task. *Journal of Sports Sciences, 30*, 1747–1756.

9.1 Introduction

To maximise sports performance, sportspeople should be in top form and have excellent equipment available. Studies in various types of sports show that interfaces (e.g., pedals, footwear, handgrip, and displays) are important determinants of success as they can help sportspeople get the most out of themselves and their equipment (Eccles, 2006; Gregor & Wheeler, 1994; Heil, Engen, & Higginson, 2004; Zadpoor & Nikooyan, 2010).

The present study focuses on car racing, one of the most physically and mentally demanding sports (Klarica, 2001). The human-equipment interface under investigation is the brake pedal, the mediating device between the driver's foot and the race car's brake system. The ability of a racing driver to control a car's braking force has a major effect on lap times. Effective deceleration of a racecar without an anti-lock braking system (ABS) requires a threshold braking (also called limit braking) technique, whereby the braking force is built up rapidly and then regulated near the optimal slip ratio of the tires (Sharp, 2009; Smith, 1996).

The brake force exerted on the wheels of a racecar approximately equals the force exerted by the driver on the brake pedal multiplied by an amplification factor (also called gain). With high gain, the pedal is sensitive and the driver needs to deliver only low pedal force to reach the required vehicle deceleration. With low gain, the pedal is less sensitive, meaning that high pedal force is required. Even though brake pedal stiffness and gain can differ by as much as a factor of 10 in racing cars (De Groot, De Winter, Mulder, & Wieringa, 2011a), little is known about the effects of gain on brake control by racecar drivers.

No research is available on the effects of gain in car racing, but there is a wealth of research on the effect of gain on operators' ability to control a visually displayed target. For an isotonic control, control-display (CD) gain is defined as the amplification between the position of the control device (e.g., pedal, joystick, knob, mouse, touchpad, trackball) and the position of the displayed output (e.g., cursor on a screen). For an isometric control (also known as a force stick), control-display gain is defined as the amplification between the applied force and the displayed output. For velocity-control (i.e., first-order) systems, control-display gain represents the amplification factor between control position/force and the velocity of the displayed output. Researchers have examined the effect of control-display gain on control performance, particularly movement time in target-acquisition tasks and tracking error in compensatory tracking tasks. Some studies have found a U-shaped profile for performance versus gain, with intermediate gains resulting in optimal performance (e.g., Accot & Zhai, 2001; Kwon, Choi, & Chung, 2011; Lin, Radwin, & Vanderheiden, 1992; MacKenzie & Riddersma, 1994). Others have reported that performance improves with gain (Bohan, McConnell, Chaparro, & Thompson, 2010; Gibbs, 1962; Huysmans, De Looze, Hoozemans, Van der Beek, & Van Dieën, 2006; Johnsgard, 1994; Thompson, McConnell, Slocum, & Bohan, 2007; Van Doorn, Unema, & Hendriks, 2005). Yet others have reported that performance decreases with gain (e.g., Andreas, Gerall, Green, & Murphy, 1957; Arnaut & Greenstein, 1986; Chase, Cullen Jr., Openshaw, Sullivan, 1965; Johnsgard, 1994; Mandryk & Gutwin, 2008; Hammerton, 1962; Schaab, Radwin, Vanderheiden, & Hansen, 1996). Casiez, Vogel, Balakrishnan, and Cockburn (2008) observed that there is "no clear result governing the effect of CD gain" (p. 222), and they themselves found that movement times in a target-acquisition task were lowest for high values of control-display gain.

One reason for the inconsistency in the literature is that control-display gain is a composite variable of control range and display range (Accot & Zhai, 2001; Arnaut & Greenstein, 1990; Buck, 1980; Ellis et al., 2004). That is, control-display gain can be increased by amplifying the displayed output while holding the required control input constant or by reducing the required control input while holding the display constant. When we restrict ourselves to studies using the former method, findings consistently indicate that increased gain results in improved performance (Andreas & Weiss, 1954; Breur, Pool, Van Paassen, & Mulder, 2010; Chase et al., 1965; Coombes, Corcos, Sprute, & Vaillancourt, 2010; Helson, 1949; Hong, Brown, & Newell, 2008; Prodoehl & Vaillancourt, 2010; Sosnoff, Valantine, & Newell, 2006; Vaillencourt, Haibach, & Newell, 2006). For example, Chase et al. (1965) let participants control the position of their extended index finger while they received feedback on an oscilloscope about errors with respect to a fixed point in space. When display amplification increased, error with respect to the target reduced, with an asymptote reached when the display range was so large that participants could not take advantage of the more precise information. Sosnoff and Newell (2006) show that tracking error can even increase when testing extremely large gains, possibly because the participants are overloaded with information that is beyond the precision of the motor system.

The more ecologically-valid means of varying control-display gain is to change the control range while holding the display constant. Buck (1980) explained: “When the display is a real life presentation as in pointing a stick or driving a car or operating a drill it may be only the control device that is susceptible to variation” (p. 580). For studies using this method of altering control-display gain, part of the inconsistency in the literature can be explained by the fact that some studies used velocity-control dynamics (Gibbs, 1962; Hammerton, 1962, Kantowitz & Elvers, 1988; Rockway, 1954; Tipton & Birmingham, 1959). Velocity-control systems have an inherent filtering property and can therefore be controlled more easily with high gain. The effect of gain in velocity-control systems crucially depends on specific design parameters, such as the maximum speed attained by the displayed output at full deflection of the control and the time required to move the control to this point (Hammerton, 1962). The results of velocity-control systems are therefore difficult to compare with the results of position-control systems.

Whether a U-curve, increase, or decrease of performance versus gain is found also depends on the range of gains under evaluation. In studies investigating a sufficiently large range of gains, a U-curve with a broad optimum appears, indicating that humans are able to achieve relatively constant performance across gains of at least a 10-fold magnitude above and below levels at which performance becomes noticeably poorer (e.g., Jellinek & Card, 1990; Jenkins & Connor, 1949; McRuer & Jex, 1967; Tipton & Birmingham, 1959). For example, Jenkins and Connor (1949) investigated gains that differed by more than a factor of 150 and found a clear U-curve of movement time versus gain. More precisely, the gains varied from .220 to 33.6 inches of pointer movement for one complete turn of a control knob. A minimum movement time was found for a gain of 1.2, while relatively similar values were obtained for gains that were three times as small or three times as large. For gains outside this range, movement times drastically deteriorated. The precise shape of the U-curve and the location of the optimum depends on numerous factors, including resolution of sensor and display, time delay, viewing distance, type of control device, backlash, inertia, friction, individual skill level, and speed versus accuracy task instructions (Ely, Thomson, & Orlansky, 1956; Helson, 1949; Gibbs, 1962; Rockway & Franks, 1959).

Helson (1949) pointed out that "the U-hypothesis expresses the fact that organisms can adapt to a fairly wide range of stimulus values and function optimally within this range" (p. 493). For tracking tasks, a broad optimum of the U-curve is in agreement with McRuer's crossover model predicting that humans adapt their own sensitivity to the gain of the controlled element (McRuer & Jex, 1967). Similarly, for target-acquisition tasks, Fitts' law (1954) predicts that movement times are unaffected when the width/amplitude ratio is constant (see also Buck, 1980; Jellinek & Card, 1990). It has been suggested that this factor (i.e., the breakdown of Fitts' law for very small targets or very large amplitudes) represents changes in movement time when gain is varied (Buck, 1980; Parng, 1988). The U-curve appears because when control-display gain is very low, factors such as maximum limb speed, fatigue, maximum input-device deflection, or the phenomenon of "clutching" (e.g., repositioning the computer mouse without affecting the display pointer, using a grip-rotate-release technique with a rotary knob, etc.) causes a performance decrease. When control-display gain is very high, limitations of fine muscle control, sensor noise, or device quantization error cause performance decrease. Quantization error arises when the maximum resolution of the control device prevents each pixel from being addressable on the display (Casiez et al., 2008).

Although the U-curve appears to be a ubiquitous phenomenon, a number of important aspects still require further evaluation. First, researchers generally used isotonic control devices, with relatively few exceptions (using isometric controls: Hess, 1973; Kantowitz & Elvers, 1988; Tipton & Birmingham, 1959). With isotonic controls, control-display gain variations produce "dramatic biomechanical changes" (Bohan, Thompson, & Samuelson, 2003, p. 442). Large movements (e.g., 1 m) tend to be carried out by the arm (shoulder & elbow joint), medium-range movements by the hand (wrist joint), and fingers are used for small movements. It is well known that the sensitivity of fingers allows for more subtle manipulations than the other body parts, and this factor may be a confounder in previous control-display gain research (Accot & Zhai, 2001).

A second issue is that most research emphasises gross performance metrics, such as average movement time and error rates (in target-acquisition tasks) or root mean-squared error (in tracking tasks). When focusing on such generic metrics little performance variation may appear between gains, even though clear differences in *control behaviour* underlie these performance outcomes. Exceptions to this limitation are the works of Accot and Zhai (2001), Casiez et al. (2008), and Mandryk and Gutwin (2008) showing that an increase of gain results in more frequent overshoot of the target or increase of maximum velocity of the displayed cursor. For target-acquisition tasks, optimum gain is reported to represent a trade-off between minimizing gross positioning time (also: slewing time, primary movement time, or ballistic movement time) on the one hand, and minimizing subsequent fine adjusting time (also: homing time, secondary movement time, steadying time, or settling time), on the other (Ely et al., 1956). Wickens (1986) provided a general hypothesis, suggesting that two counteracting effects play a role in determining optimum control-display gain: stability and workload.

In this study we examined the effect of gain between brake pedal force and visually displayed output (i.e., control-display gain) on the ability of racing car drivers to track a visually displayed target value. The study used an isometric brake pedal, that is, a pedal with a force sensor and no movement, thereby avoiding clutching and effects of limb speed. We chose a broad range of control-display gains, from high (i.e., a highly sensitive pedal) to low, so that the participants were physically just able to deliver the pedal force required to carry out the task successfully. We made a distinction between a step-function (i.e., target-acquisition) task and multisine (i.e., random input-tracking) tasks. The step-function tasks allowed us to investigate build-up speed as well as overshoot, whereas the multisine task allowed us to investigate how participants behave under continuous control. We are not aware of any previously published research that makes the same distinction, although Wickens (1986) has addressed the importance of introducing both kinds of paradigm to acquire a complete picture of the effect of gain. Based on Wickens' framework, we hypothesised that a low control-display gain would result in improved stability, namely fewer overshoots and increased consistency (i.e., lower within-participant variability) compared to high control-display gain. Furthermore, we hypothesised that because lower control-display gain requires larger forces to be delivered, this would be manifested by a slower build-up of the displayed output value and higher subjective physical workload.

9.2 Method*!*

9.2.1 Participants

The research was approved by the TU Delft Human Research Ethics Committee and all participants provided written informed consent. The participants were 13 males with a mean age of 22.6 years (*SD* = 5.9 years). The participants were drivers with experience in European racing classes, including Formula 3 (2 participants), Formula 2 (2 participants), Formula Ford (2 participants), Formula Renault 2.0 (1 participant), World Series by Renault (1 participant), Radical Cup (1 participant), as well as Formula Student racing (2 participants). The two authors of this research also took part in the study. The first author has obtained a club racing license and the second author has been involved in Formula 3 racing. The mean self-reported number of races per year and tests per year, excluding the two authors' data, was 15.9 (*SD* = 9.4) and 24.9 (*SD* = 20.1), respectively. Participants received a token financial remuneration of 10 euro.

9.2.2 Apparatus

For the experiment, a formula racing cockpit was fitted with a 19-inch CRT monitor (Compaq S9500; Figure 1). We used a CRT monitor because it offers minimal time delay between control input and displayed output (which we estimated at 0.05 s) as compared to modern television screens and flat screen monitors (estimated at 0.12 s). The resolution of the screen was 1,280 x 1,024 and the refresh rate was 85 Hz. The brake pedal was immobilised, and the brake pedal force was measured using a strain gauge. The pedal surface area was 90 x 40 mm rectangular. The measured force was recorded at 60 Hz.

9.2.3 Display

A target signal and an output signal were displayed side by side on the monitor (Figure 2). After each braking action (henceforth: trial), the participants were shown a score between 1 and 10 at the top of the display, indicating how well they had tracked the target value, 10 meaning a perfect score. This score was a linear function of the root mean-squared error between output value and target value. The participants were shown their best score in the session between parentheses. Pre-

Figure 1. Photo of the experimental setup. All participants used headphones to block out environmental sounds.

vious research indicates that knowledge-of-results (KR) feedback enhances motivation and results in increased performance as compared to receiving no KR feedback (e.g., Locke, 1968). To allow participants to anticipate the target, two triangles at the top of the display gradually approached each other from both sides and hit each other precisely at the moment of onset of the target value. Figure 2 shows these two triangles after onset and prior to onset.

9.2.4 Procedures

The participants were first informed in writing about the aim of the experiment, about the functionality of the display, and about their task to follow the displayed target value as precisely as they could. The instructions further stated that participants were free to use their right or left foot, but that they were not allowed to swap feet during the experiment. Participants were allowed to choose which foot they used, because some racecar drivers are accustomed to braking with their left foot while others are accustomed to using their right foot. All participants completed the experiment with racing shoes or shoes with soles similar to racing shoes. Participants were informed

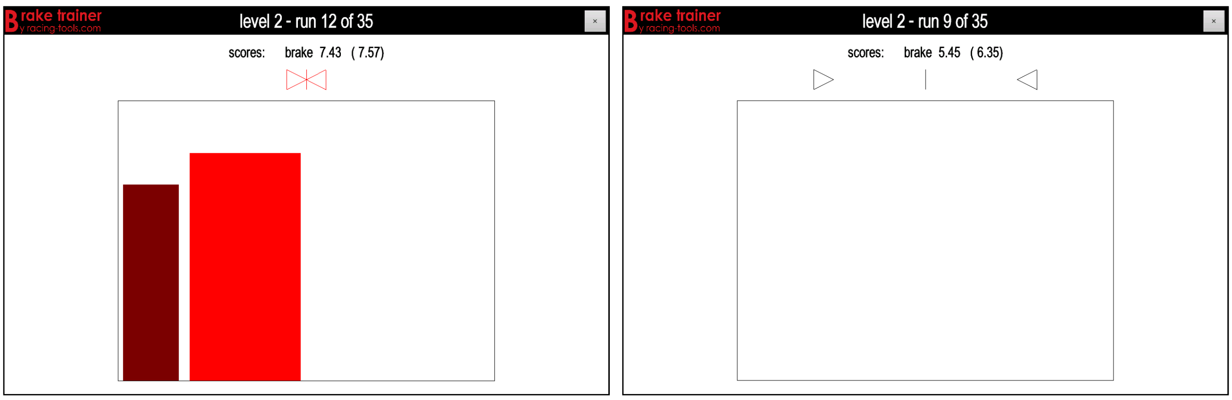

Figure 2. The display used in the experiment. Left: The dark bar represents the target value and the lighter bar represents the delivered brake pedal force. The visible scale ranged from 0 to 100%. The triangles were approaching each other and hit exactly when the target value onset changed from 0 to 70% (at time = 0.5 s). Right: the display prior to onset of the target value; the two triangles were approaching each other from both sides of the screen.

that they were allowed to ask questions or abort the experiment at any time they wished.

After reading the instructions and providing written informed consent, participants sat in the cockpit and put on regular headphones to block out ambient noise. No audio signal was provided. Each participant completed eight sessions: four step-function sessions with different control-display gains per session followed by four multisine sessions with the same control-display gains. For both step-function and multisine sessions, the four control-display gains were randomised, also ensuring that the lowest and highest gain did not follow each other. The control-display gains were such that the pedal forces required to achieve a 70% output value on the display were 100 N, 400 N, 700 N, and 1,000 N, respectively.

Each step-function session consisted of 35 trials of 2 s step-function tasks, in which the target value instantaneously rose from 0 to 70% (see Figure 3). There was a 4 s pause between each trial. Each multisine session consisted of 15 trials of 18 s quasi-random signals as a target value. The quasi-random signals were the sum of five sinusoids with frequencies of 1 Hz, 0.5 Hz, 0.2 Hz, 0.1 Hz, and 0.05 Hz and random phase shifts. The amplitude of the target value varied between 2 and 33% (see Figure 4). There was a 5 s pause between trials.

Following each of the eight sessions, participants completed the NASA-Task Load Index (TLX) to assess mental and physical workload. This questionnaire has been used in at least 550 human-factors studies, and its psychometric properties are well documented (e.g., Hart & Staveland, 1988; Hart, 2006; Nygren, 1991; Rubio, Díaz, Martín, & Puente, 2004).

9.2.5 Dependent measures

First, the logged data (time vector, target value, and output value) were interpolated to 1,000 Hz. Next, the following dependent measures were calculated for the step-function task per control-display gain and per participant.

Error (%). Root Mean-Squared Error (RMSE) between the target and output signal between t = 0 and 2.0 s, averaged over trials 11–35.

BuildupSpeed (%/s). The build-up speed was calculated as 50% divided over the time taken to rise from 10 to 60%. For example, if a participant took 0.2 s to rise from 10% to 60%, then the build-up speed of that trial was 250%/s. BuildupSpeed was averaged over trials 11–35. The 10 and 60% cut-offs were based on pilot measurements and visual inspection of the results. 10% is required to reach maximum speed, whereas 60% is 10% below the target value of 70%.

PeakValue (%). The maximum value of the output value between t = 0 and 2.0 s, averaged over trials 11–35.

Variability (%). The standard deviation per time point between t = 0 and 2.0 s, averaged over the time points and over trials 11–35. This measure represents the within-participant variability.

The following dependent measures were calculated for the multisine task per control-display gain per participant.

Error (%). Root Mean-Squared Error between the target and output signal between t = 0 and 18.0 s, averaged over trials 4–15.

Variability (%). The standard deviation per time point between t = 0 and 18.0 s, averaged over the time points and over trials 4–15. This measure represents the within-participant variability.

For both step-function and multisine tasks, the TLX items (Mental demand, Physical demand, Temporal demand, Performance, Effort, and Frustration) served as dependent measures. The responses were expressed on a scale from 0% (very low/perfect) to 100% (very high/failure).

Note that the first trials per session were excluded to prevent adaptive effects between gains (cf. O'Dwyer & Neilson, 1996). The cut-off values (trial 11 for the step function and trial 4 for the multisine) were based on a pilot experiment with different drivers.

9.2.6 Statistical analyses

For each dependent measure, the effects of gain were analyzed by submitting the data matrix (13 participants x 4 gains) to the Friedman test, which is the nonparametric version of two-way analysis of variance (ANOVA). A nonparametric test was used because of the heteroscedasticity of the dependent measures for the different gain levels. Whether the participation of the two authors acted as a source of bias (e.g., due to their knowledge of the experiment setup and hypotheses) was investigated by repeating the analyses with the data of the authors removed.

9.3 Results

Tables 1 and 2 show the results per dependent measure. The Error measures for the two tasks display a slight U-curve pattern, with the lowest and highest gains performing less accurately than the middle two gains. The effect sizes between gains were moderate, yielding significant effects for the step function (p = .013) and marginal effects for the multisine function (p = .078). A two-sided Wilcoxon signed ranks test between pairs of gains revealed a significant difference between the highest gain (100 N) and the second lowest gain (700 N), p = .005 for the step-function task, and p = .033 for the multisine task.

The other dependent measures revealed larger differences between gains, having a monotonic trend. Lower control-display gains consistently resulted in lower build-up speeds, lower maximum values (i.e., lower overshoot), and more consistent performance within participants. Furthermore, the lower the gain, the higher the subjectively reported Physical demand, which was a large and significant effect. Differences between gains were found to a lesser extent for Effort. Self-reported Mental demand, Temporal demand, Performance, and Frustration revealed no significant differences between gains.

Table 1. Means (with standard deviations in parentheses) of the objective dependent measures per gain, expressed in pedal force in Newtons required to deliver an output value of 70% (1,000 N = lowest gain, 100 N = highest gain).

Variable	1,000 N	700 N	400 N	100 N	χ^2 (3,51)	W	p
Step-function task							
Error (%)	8.62 (1.03)	8.33 (1.68)	8.60 (1.37)	9.21 (1.80)	10.8	.276	.013
BuildupSpeed (%/s)	364 (103)	505 (292)	516 (181)	650 (221)	21.2	.553	.000
PeakValue (%)	72.5 (1.6)	74.5 (1.9)	75.9 (2.3)	78.3 (2.9)	33.7	.863	.000
Variability (%)	3.04 (0.44)	3.19 (0.85)	3.62 (0.62)	4.84 (1.12)	29.0	.744	.000
Multisine task							
Error (%)	2.31 (0.45)	2.24 (0.37)	2.32 (0.50)	2.47 (0.50)	6.8	.174	.078
Variability (%)	1.07 (0.41)	1.03 (0.40)	1.12 (0.30)	1.42 (0.35)	15.7	.401	.001

W is Kendall's *W*, also known as Kendall's coefficient of concordance. If W = 0, then there is no correspondence amongst the participants and the results may be regarded as essentially random. If *W* is 1, then all participants have the same rank order of gains.

Table 2. Means (with standard deviations in parentheses) of the self-reported dependent measures per gain, expressed in pedal force in Newtons required to deliver an output value of 70% (1,000 N = lowest gain, 100 N = highest gain).

Variable	1,000 N	700 N	400 N	100 N	χ^2 (3,51)	*W*	*p*
Step-function task							
Mental demand (%)	46 (22)	50 (21)	46 (25)	48 (29)	1.1	.027	.787
Physical demand (%)	73 (18)	52 (16)	31 (24)	16 (20)	29.5	.757	.000
Temporal demand (%)	34 (19)	33 (19)	27 (17)	30 (25)	4.2	.107	.243
Performance (%)	38 (11)	45 (15)	41 (12)	47 (23)	2.7	.069	.445
Effort (%)	71 (8)	57 (19)	43 (21)	44 (25)	15.9	.407	.001
Frustration (%)	26 (22)	31 (17)	30 (21)	33 (25)	1.3	.033	.728
Multisine task							
Mental demand (%)	63 (14)	60 (14)	55 (17)	58 (20)	3.3	.083	.355
Physical demand (%)	55 (18)	42 (23)	28 (23)	23 (28)	30.4	.779	.000
Temporal demand (%)	37 (21)	30 (19)	34 (24)	29 (20)	4.0	.101	.266
Performance (%)	37 (14)	35 (16)	38 (15)	42 (16)	1.8	.047	.606
Effort (%)	65 (17)	58 (19)	55 (23)	58 (22)	7.6	.196	.054
Frustration (%)	34 (21)	27 (21)	26 (21)	38 (28)	4.7	.119	.199

W is Kendall's *W*, also known as Kendall's coefficient of concordance. If $W = 0$, then there is no correspondence amongst the participants and the results may be regarded as essentially random. If *W* is 1, then all participants have the same rank order of gains.

Figures 3 and 4 illustrate the averages of control outputs per time point for the step-function and multisine tasks, respectively. Qualitatively, there are only small differences between control-display gains. However, within-participants variability (Figures 5 and 6) reveals larger differences between gains, with the highest gain (100 N pedal force for a 70% output value) yielding the lowest consistency within participants (see also Table 1).

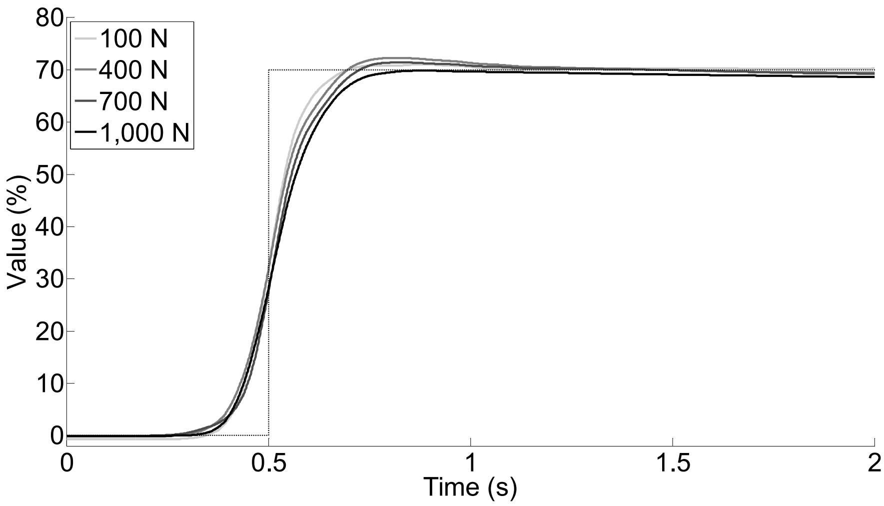

Figure 3. Average outputs in the step-function task. The dotted line represents the target value (always the same; independent of gain). The solid black and gray lines represent the delivered values averaged per time point over trials 11–35, averaged over all participants.

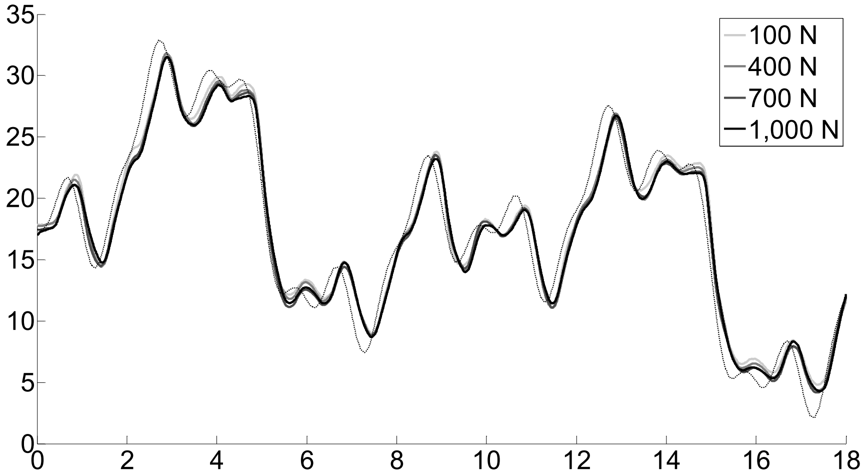

Figure 4. Average outputs in the multisine tasks. The dotted line represents the target value (always the same; independent of gain). The solid black and gray lines represent the delivered values averaged per time point over trials 4–15, averaged over all participants.

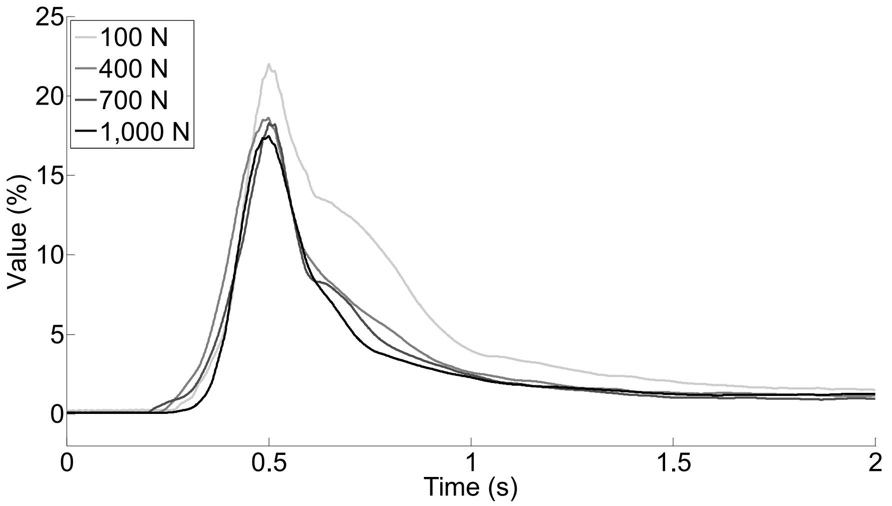

Figure 5. Within-participants variability in the step-function task. The black and gray lines represent the within-participant standard deviations per time point for trials 11–35. The standard deviations were subsequently averaged over all the participants.

Figures 7 and 8 illustrate that the adaptation effect between gains was restricted to the first few trials per gain, justifying our decision to exclude the first 10 trials in the step function task. Figure 9 illustrates the results for one participant with the delivered pedal force on the y-axis. It can be seen that although the highest gain resulted in the most consistent performance in Newtons, it result in the lowest consistency *relative* to the target value (cf. Figure 5).

Finally, the Friedman tests were repeated with the two authors' data removed. For the six dependent measures reported in Table 1, the *p* values were .014, .000, .000, .000, .241, and .001, while the corresponding Kendall *W*'s effect sizes were .322, .557, .841, .769, 127, and 517. These *p* and *W* values are in close agreement with the *p* and *W* values for all 13 participants reported in Table 1, which indicates that no bias was introduced by the authors' participation.

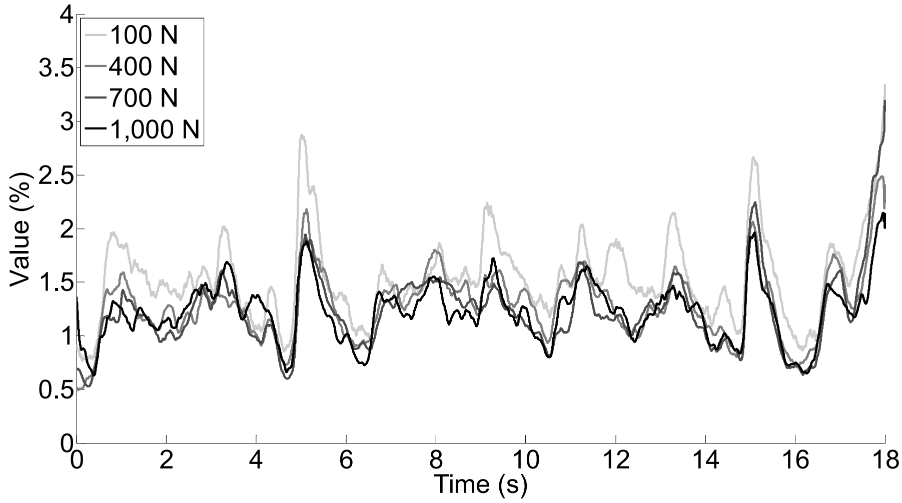

Figure 6. Within-participant variability in the multisine task. The black and gray lines represent the within-participant standard deviations per time point for trials 4–15. The standard deviations were subsequently averaged over all the participants.

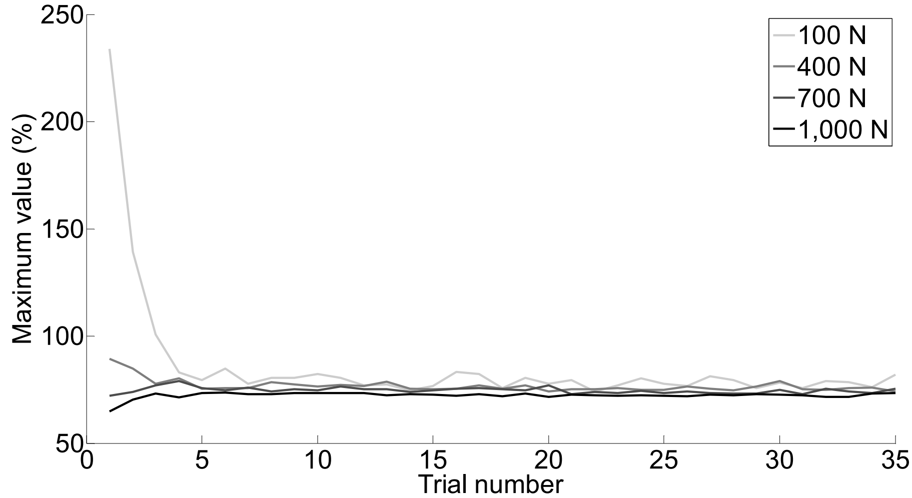

Figure 7. Maximum value (%) as a function of trial number in the step-function task. The black and gray lines represent the averages over all participants per trial.

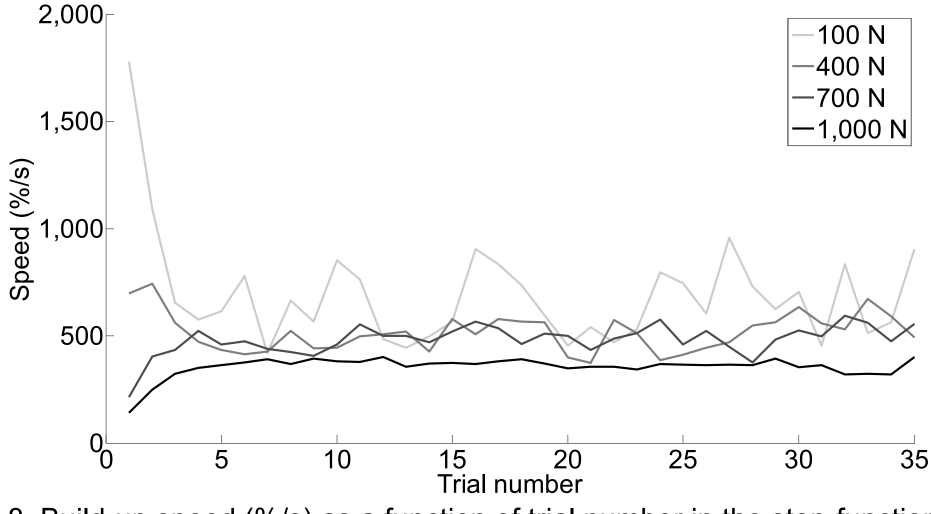

Figure 8. Build-up speed (%/s) as a function of trial number in the step-function task. The black and gray lines represent the averages over all participants per trial.

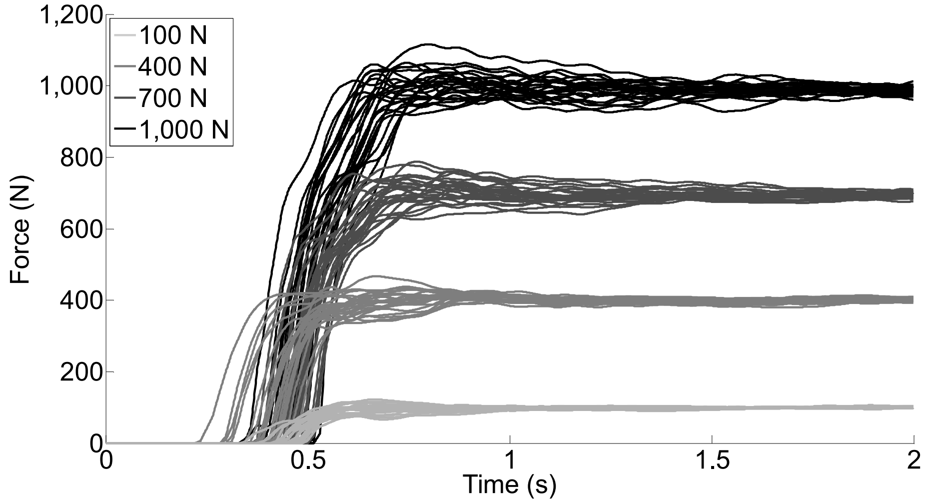

Figure 9. Delivered force for one representative participant for the four gain levels in the step-function task.

9.4 Discussion

This study investigated the effect of control-display gain using an isometric brake pedal. Overall, the hypotheses were confirmed. The Error metrics and average outputs per time point revealed a broad optimum with small differences between gains. However, clear effects were found for within-participants variability, maximum speed, overshoot, and subjective physical load. These results provide support for Wickens' (1986) hypothesis that choosing an optimum gain involves balancing stability against physical effort. Higher gain results in higher speed of the output value, but this advantage is offset by inconsistent control and more frequent overshoot.

The results of the NASA-TLX showed that control-display gain had a strong effect on Physical demand and a relatively minor effect on Effort, the latter term being semantically related to Physical demand. No significant effects were found for the four remaining TLX items (Mental demand, Temporal demand, Performance, & Frustration), which were cognitively rather than physically oriented. The TLX results can be explained by the fact that the visual target presented to the participants was exactly the same for the four gain levels, while the physical forces that had to be delivered varied, with the lowest gain requiring a 10-time larger force than the highest gain. The constant Mental demand as a function of gain can be interpreted as workload-performance dissociation (Hancock, 1996; Yeh & Wickens, 1988). That is, while BuildupSpeed, PeakValue, Variability, and to a lesser extent Error, were all affected by gain level, Mental demand always stayed at the same level. This workload-performance dissociation may indicate that the participants were fully committed to the task, being insensitive to their own output (Hancock, 1996). Our study attempted to fully engage participants in the task by providing knowledge-of-results (KR) feedback after each trial. Pilot experiments without such feedback revealed that participants reported high levels of boredom and difficulty concentrating, and we believe that this may also have played a role in much of the previously published research on the effect of control-display gain (cf. Kantowitz & Elvers, 1988, for an exception using KR feedback).

In our study, we sought to find an effective compromise between realism and experimental control. Evans (2004) explained that "any decisions regarding major investments in additional driver simulators should identify what specific problems they can be used to solve, and why they can solve them when only slightly less sophisticated simulators could not" (p. 190). Some aspects of our experiment were realistic, such as real racing drivers sat in a real cockpit. Other aspects were not simulated at all: no racetrack was projected in front of the participants, participants did not have to steer, the cockpit did not move, and there were no vibrations (see Lewis & Griffin, 1977 for a study on control-display gain and vibrations). Yet other aspects were abstractions of a real racing task. For example, we used an isometric pedal whereas pedals in racecars have some (limited) displacement, and the step function roughly resembles the threshold braking manoeuvres that racecar drivers actually use. We used a simple simulator to strip the research question down to its essentials and maximise controllability and repeatability. In pilot experiments, we evaluated the effect of brake pedal parameters on a simulated racetrack, and while it was possible to keep tire wear and environmental conditions under control in virtual simulation, there were still a large number of uncontrolled parameters (e.g., entry speeds, braking points, different driving styles) which hampered experimental control. It can also be questioned whether a high-fidelity simulator, such as a motion-base simulator, would be desirable. Because of kinematic and dynamic constraints of motion platforms, no simulator can replicate all the acceleration forces that are produced in a real racecar. No motion may therefore be better than bad motion (cf., Bürki-Cohen, Soja, & Longridge, 1998), questioning the value of motion platforms for flight training.

Further research could investigate whether the findings can be generalised to a real racing environment. The question is how comparable the relatively static design of this study is to the far more dynamic racing environment, where drivers experience brake forces not only visually but also directly, through pressure receptors and the vestibular organs. However, it should be noted that renting a racetrack is costly, and that it is mechanically challenging to change the brake system of a racecar in a time-efficient manner. Furthermore, track conditions, temperature, humidity, and tire wear affect braking performance in subtle but important ways, making it extremely difficult to obtain a scientifically meaningful result.

The findings of the present study add to our understanding of control-display gain, an important parameter in all continuous control devices. Our results showed that there is a broad range of gains at which participants reach optimal *average* performance. However, large effects of gain were found regarding the variability of performance, and physical workload. Our findings offer prospects for improving human-equipment interfaces in sports, and for developing brake pedals for racing cars in particular. For example, if consistent performance is required and outliers should be avoided (e.g., on wet roads or in cases where tasks are difficult), then a low gain may be preferred. However, if fast build-up speed of brake pressure is preferred, or if drivers are physically tired, then higher gain may be needed.

Conclusions and recommendations

This chapter summarizes the main conclusions and recommendations of this thesis. The text is structured along the three parts of the thesis.

Part 1. **Driver performance** in fixed-base driving simulators

Chapter 1. Braking to a full stop at a prescribed target position on the road is a driving manoeuver regularly used in experiments to investigate driving behavior, or to test vehicle-acceleration feedback systems in simulators. Many different performance measures have been reported in literature. A computer simulation and empirical driving simulator experiment were conducted to investigate various performance measures. Reliable and valid measures were found to be (a) the speed at braking onset, (b) the distance to the target position at braking onset, (c) the stopping position relative to the target position, and (d) a novel metric that measures the deviation from constant deceleration.

Chapter 2. The effectiveness of a driving simulator depends to a large extent on its level of realism, also called fidelity. Simulator motion platforms can be used to provide physical motion cues in driving simulators, which often leads to improved fidelity, but at significant financial cost. In this chapter eight low-cost alternatives for providing motion cues were tested by comparing driver performance with these cues to a control group that drove without the system. Overall, the eight non-vestibular cueing systems showed desired effects: drivers adopted lower speeds and had smaller decelerations while braking. Also, the subjective realism score of the simulation tended to improve. The study shows that providing acceleration feedback to drivers can improve realism of driving performance, irrespective of whether the acceleration cues are physically realistic. However, the driver integrates the available cues in an adaptive fashion, and the visual, tactile, or vestibular cues available in a simulator are different than the cues that are available in a real car. A potential disadvantage of using ‘unrealistic’ cues for driver training may be that the transfer to the real vehicle is impaired. It is recommended to conduct transfer-of-training studies to investigate how performance learned in the simulator transfers to on-road driving.

Chapter 3. Most instructions given during driver training in current simulators are presented through speech, that is, in the auditory modality. This approach corresponds to regular on-road driver training where a human driving instructor is available, and is furthermore believed to interfere little with the predominantly visual driving task. This chapter investigated the effect of presenting instructions to the trainee in the visual modality. It is experimentally shown that visual and multimodal (combined visual and speech) route instructions yield improved adherence to the instructions than speech route instructions. Most participants preferred the visual route instructions. Because large individual differences exist, it is suggested to investigate the effectiveness of adaptive instruction systems that can vary the modality to the need of the individual student.

Chapter 4. Apart from the need to satisfy simulator fidelity requirements, more attention is needed to the didactical properties of a simulator training curriculum. This chapter evaluated the didactical properties of various driver training simulators in the Netherlands. It is shown that the first principles of instruction (Merill, 2002a, 2002b)—(1) problem-centered, (2) activation, (3) demonstration, (4) application, and (5) integration—are not implemented to their full potential. Modern techniques provide opportunities to implement each of these principles in simulator-based driver training, to improve the didactical quality of training. It is recommended to empirically investigate the effect of training methods, and to validate the use of the first principles of instruction in driver training.

Chapter 5. Augmented feedback (feedback other than the naturally available task-intrinsic feedback) is a new trend in simulator-based driver training of tasks such as lane keeping and car following. For effective training the augmented feedback should be provided sparingly and in such a way that the learner does not become dependent on it. This chapter investigated bandwidth feedback, which is feedback that depends on whether performance is within or outside a pre-set performance limit, in the task of lane-keeping. Previous motor learning research shows that bandwidth feedback improves learning, and that off-target feedback (feedback only when performance is outside the pre-set limits) yields better transfer performance than on-target feedback (feedback only when performance is within the pre-set limit). A between-subjects experiment was conducted with four groups: 1) on-target feedback of seat vibrations when close to the lane center, 2) off-target feedback, 3) no feedback, and 4) 'realistic' vibrations. After practice, subjects had an immediate and a one-day delayed retention session with the realistic settings only. During practice, both the on-target and off-target groups outperformed the non-augmented groups, supporting the findings in Chapter 2 which found that, on average, drivers are able to use any additional cue, realistic or unrealistic, to improve their performance. The difference in performance between the augmented and non-augmented groups diminished, however, in the retention phase, with still a small performance benefit for the group who trained with the off-target concurrent feedback. It was concluded that for the lane-keeping task off-target feedback is superior to on-target feedback, in line with previous findings but now also validated for an ecologically valid task. For effective learning of tracking skills, the onset of a stimulus should be associated with erroneous performance, not with correct performance. It is recommended, however, to study the longer-term effects of augmented feedback, e.g., in more delayed retention tasks, or in realistic (and more complex) driving environments.

Chapter 6. Previous research in learning motor-tasks has shown that degrading the task conditions during practice (that is, making the task more difficult to perform) can enhance longer-term retention performance. This chapter investigated the effects of the tire-road friction coefficient on learning a self-paced lane-keeping task in a driving simulator. We hypothesized that: 1) practicing with low-grip (LG) tires would result in better lane-keeping performance and lower driving speeds during retention with normal-grip (NG) tires, than practicing with NG tires, and 2) practicing with high-grip (HG) tires would have the opposite effect. Results of an experiment confirmed the hypothesized effects on driving speed, but not those on lane-keeping performance which showed mixed results. Even though varying the tire grip had a considerable effect on vehicle dynamics, the subjective workload reported by participants was unaffected. Tentatively, in the more demanding LG condition participants managed their workload level by driving less fast, and in the less demanding HG condition, participants managed their workload by improving their task performance, that is, accurately adhering to the lane center and reducing off-nominal events. The results clearly show that drivers have many different behavioral mechanisms available to compensate for changing task settings and task demands. It is concluded that driver training effectiveness in principle can be enhanced by providing the drivers with low-grip tires, but one should be aware that drivers can compensate for this by changing their behavior such that it satisfies their own intrinsic values. Our results suggest that it is effective to increase the task demands during practice and allow trainees to explore the limits of tolerable behavior for themselves by making errors at their own pace. The driver simulator is suited for this kind of training. It is recommended to study the possible wider applications of this finding for other driving tasks. In addition, transfer of training studies should be conducted to assess what aspects of driving skill and driving style found in the driving simulator can be generalized to the operational environment, and whether the simulator training can help to improve safe driving on the roads.

Chapter 7. In this chapter we performed a second study to investigate whether simulator-based driver training becomes more effective when trainees are induced to make more errors, and learn from their occurrences. In a racecar driving setting, we intentionally degraded the handling qualities of the vehicle to investigate that effect on lap time performance in retention. Again the tire-road friction coefficient was manipulated between low-grip (LG), normal (NG), and high-grip (HG). A between-subject experiment was conducted where inexperienced racecar drivers completed four 10-minute practice runs on an oval track, driving a Formula 3 car, each group with one grip setting. After practice, two retention sessions followed: one with NG in a Formula 3 car, and another session with a Formula 1 car. Results show that whereas lap times with the LG were significantly lower than HG in the first retention session, driver confidence was higher and the level of frustration was lower. Apparently, the low-grip group became more complacent with their performance, drove at a more comfortable speed, and was less inclined to explore the grip limits of the racecar when the vehicle handling improved.

Part 3. **Racing simulator** validity and controllability

Chapter 8. Car racing is a very demanding activity with limited possibilities to train drivers and test car equipment and settings. Racecar simulators offer potential in providing opportunities for safe, efficient and standardized training and testing. In this chapter we report the results of a racecar simulator validation study, in which we compared the fastest lap times of a small sample of drivers during training sessions in the simulator with their fastest lap times during real-world race events. The correlation was found to be statistically significant, showing the predictive value of measures obtained in the simulator. In addition, we investigated the effects of brake pedal stiffness on driver performance in the simulator, in two levels of task difficulty. Section times were not significantly affected, which again indicates the remarkable adaptive capability of the human driver to compensate for changing task settings. It was concluded that, because of its predictive capabilities and despite the relatively low-fidelity of the cueing environment which may still be improved, the racing simulator is a viable alternative for on-track testing. It is recommended to investigate the possible effects of simulator fidelity on learning and adaptation.

Chapter 9. To minimize lap times during car racing it is important for drivers to exhibit rapid and precise brake forces on the vehicle. In this chapter the effects of the gain (that is, the control-display amplification factor) between the brake pedal force and a visually presented output signal are investigated, in a task that required drivers to accurately track a target. The test set-up was a formula racing car cockpit fitted with an isometric (i.e., force control, no position changes) brake pedal. Thirteen racing drivers performed tracking tasks with four control-display gains and two target functions, a step and a multi-sine signal. Results show that although the control-display gain had only a minor effect on the tracking performance, it considerably affected the build-up speed, overshoot, within-subjects variability, and self-reported physical load. These results provide evidence for Wickens' (1986) hypothesis that choosing an optimal gain means to establish a balance between control loop stability and physical effort. That is, higher gains yield higher build-up speeds of the controlled variable, but this benefit is partly lost by more inconsistent control and overshoot. It is recommended to investigate whether these results can be generalized to a real racing environment, which is much more dynamic and immerses the driver in a much richer cueing environment.

Conclusions

From the above, the main conclusions from this thesis can be stated as follows:

- Low-cost non-vestibular motion feedback systems can positively influence driving performance in a fixed-base driving simulator (Part 1, Chapter 2)
- Presenting instructions in the visual modality rather than the commonly used auditory modality can improve adherence to the route instructions (Part 1, Chapter 3).
- The didactical possibilities of current driving simulators are not fully exploited (Part 2, Chapter 4).
- Simulators allow for training programs that are (practically, financially, ethically) challenging in a real-world environment because, in a simulator, novice drivers can make errors and learn from those errors systematically, in their own pace, without physical risk (Part 2, Chapters 5, 6, and 7).
- Concurrent off-target augmented feedback is beneficial for learning the lane keeping task and is superior to on-target augmented feedback (Part 2, Chapter 5).
- It is effective to increase the task demands during practice and allow trainees to make errors at their own pace (Part 2, Chapter 6).
- Drivers have the ability to adapt to changing task settings, presumably to satisfy the performance and/or mental workload levels. (Part 1, Chapter 2; Part 2, Chapters 5, 6 and 7; Part 3, Chapter 8).
- Choosing an optimal control-display gain in tracking tasks requires the designer to strike a balance between stability and physical effort (Part 3, Chapter 9).

Recommendations

Based on this thesis, the following recommendations can be given for future research:

1. To further advance the state-of-the art in driving simulator training it is important to conduct **transfer-of-training** studies. Specifically:
    a. The effects of using ‘unrealistic’ performance-enhancing cues in training simulators on performance in real driving tasks should be investigated. Such studies may clarify whether these unrealistic cues enhance or degrade driving performance in a real vehicle (Part 1, Chapter 2).
    b. If the “first principles of instruction” would be implemented in driving simulators, would this lead to improved driving skill and driving style on the roads? (Part 2, Chapter 4)
    c. What aspects of driving skill and driving style, learned in the driving simulator, can be generalized to the operational environment, and can simulator training help to reduce unsafe driving on the roads? (Part 2, Chapter 6)

2. This thesis conducted several studies that touched upon the required level of **fidelity** of a driving simulator for learning a driving task. This thesis contains several recommendations to further study the effects of fidelity, a selection of which are listed below:
    a. Would low-cost non-vestibular cueing systems in low- to medium-fidelity simulators yield better or worse performance and transfer-of-training effects as real physical motion systems in high-fidelity simulators? (Part 1, Chapter 2)
    b. It is recommended to study the validity of some findings in this thesis in more high-fidelity simulators, preferably even real driving, for example, the effects of augmented feedback on retention performance (Part 2, Chapter 5), and the investigations discussed in the two chapters on racecar simulators (Part 3).
3. Regarding the design of a simulator-based training program, it is recommended to investigate the possibility of providing an adaptive and learner-centered curriculum, including for instance the use of trainee-adaptive modalities in communication (Part 1, Chapters 1 and 3).

In conclusion, this thesis explored various ways to exploit driving-simulator software and hardware systems to increase the effectiveness of driving simulator training. Only a selection of the many possible ways to improve simulator training has been investigated, and current-day simulator designers have a massive freedom in selecting and tuning their cueing apparatus. Yet, relatively little is known about how these settings affect the effectiveness of the training, many issues still need to be resolved, and a large part of the possible cueing settings and combinations of systems remain largely unexplored. One of the main findings is that, whatever combination of cueing systems and devices is used, humans are effective in adapting their behavior to the set of cues made available to them to improve their performance and/or keep their workload at an acceptable level. One may need to delve deeper into the human neurology and physiology to investigate how humans process and adapt to the available information. Psychophysiological measurements (e.g., heart rate, eye-tracking, galvanic skin response, and electromyography) may be useful research tools for this purpose. In addition, functional and structural brain imaging techniques could help clarify within- and between-subject differences in information processing and decision making. Cognitive and cybernetic modeling techniques could be useful for understanding a driver from a control-theoretic perspective.

The knowledge gained in this thesis may be of direct use or at least practical relevance for practitioners in simulator-based driver training. Partly based on this thesis project, I initiated a simulator center for the training of racecar drivers in Delft (Figure 1).

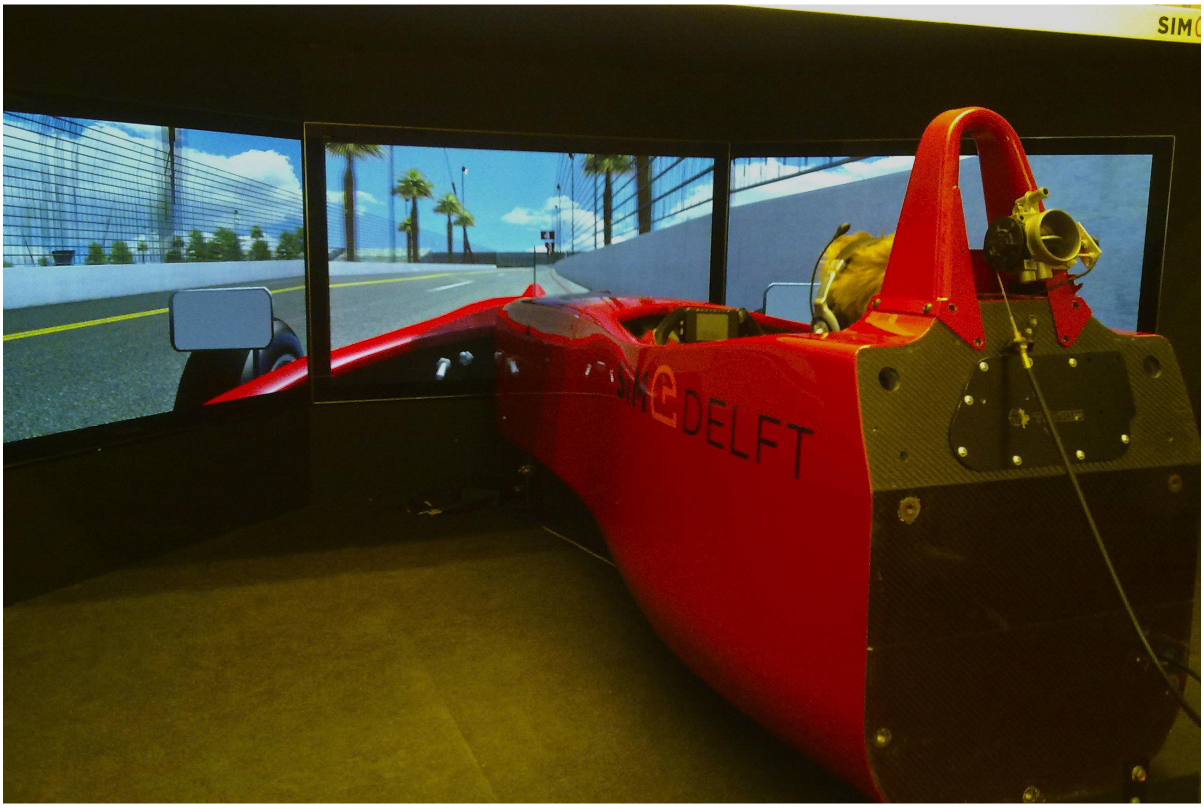

Figure 1. Fixed-base racing simulator as used for racecar driver training in Delft.

References

Aarts, L., & Van Schagen, I. (2006). Driving speed and the risk of road crashes: A review. *Accident Analysis & Prevention, 38*, 215–224.

Accot, J., & Zhai, S. (2001). Scale effects in steering law tasks. *Proc. CHI 2001: ACM Conference on Human Factors in Computing Systems*, 1–8.

Ackerman, P. L. (1987). Individual differences in skill learning: An integration of psychometric and information processing perspectives. *Psychological Bulletin, 102*, 3–27.

Adams, J. A. (1964). Motor skills. *Annual Review of Psychology, 15*, 181–202.

Advisory Group for Aerospace Research and Development (1980). *Fidelity of simulation for pilot training* (AGARD-AR-159). Neuilly sur Seine, France: North Atlantic Treaty Organization.

Allen, R. W., Park, G. D., Cook, M. L., & Fiorentino, D. (2009). Training and assessment of novice drivers. *Proceedings of the Driving Simulation Conference Europe*, Monte Carlo, Monaco, 91–102.

Allen, R. W., Park, G. D., Cook, M. L., & Fiorentino, D. (2007). The effect of driving simulator fidelity on training effectiveness. *Proceedings of the Driving Simulation Conference North America*, Iowa City, IA.

Allen, R. W., Rosenthal, T. J., Anderson, F., Aponso, B. L., Harmsen, A., & Markham, S. (1999, July). Low cost, PC based techniques for driving simulation implementation. *Paper presented at the Driving Simulation Conference*, Paris, France. Retrieved from http://www.systemstech.com/index.php?option=com_docman&task=doc_download&gid=34&Itemid=72

Andreas, B. G., Gerall, A. A., Green, R. F., & Murphy, D. P. (1957). Performance in following tracking as a function of the sensitivity of the airplane-type control stick. *The Journal of Psychology, 43*, 169–179.

Andreas, B. G., & Weiss, B. (1954). *Review of research on perceptual-motor performance under varied display-control relationships* (Scientific Report No. 2, AF 30 (602) – 200). Rochester, NY: The University of Rochester.

Arnaut, L. Y., & Greenstein, J. S. (1986). Optimizing the touch tablet: The effects of control-display gain and method of cursor control *Human Factors, 28*, 717–726.

Arnaut, L. Y., & Greenstein, J. S. (1990). Is display/control gain a useful metric for optimizing an interface? *Human Factors, 32*, 651–663.

Ashworth, B. R., McKissick, B. T., & Parrish, R. V. (1984). *Effects of motion base and g-seat cueing on simulator pilot performance* (NASA Technical Paper 2247). Washington, DC: National Aeronautics and Space Administration.

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*, 145–169.

Barnard, Y.F., 2006, Didactical and pedagogical aspects of e-learning tools. In *Proceedings of the conference on European guidelines for the application of new technologies for driver training and education*, J.P. Pardillo (Ed.) (Madrid: Universidad Politécnica de Madrid).

Beckman, B. (1991). *The physics of racing*, Part 5: Introduction to the racing line. Available online: http://www.miata.net/sport/Physics/05-Cornering.html

Bédard, M. B., 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*, 336–340.

Bekiaris, E. (2007, December). Modules for the simulators of the future. Paper presented at the TRAIN-ALL 1st pan-European Workshop: Integrated and Interoperable Training Simulators of the Future, Sulzbach-Rosenberg, Germany. Retrieved from http://www.trainall-eu.org/documents/HIT2_Presentation.pdf

Bella, F. (2008). Driving simulator for speed research on two-lane rural roads. *Accident Analysis & Prevention, 40*, 1078–1087.

Bilodeau, E. A., & Bilodeau, I. M. (1961). Motor-skills learning. *Annual Review of Psychology, 12*, 243–280.

Boer, E. R., Girshik, A. R., Yamamura, T., & Kuge, N. (2000). Experiencing the same road twice: A driver centered comparison between simulation and reality. *Proceedings of the Driving Simulation Conference Europe*, Paris, France.

Boer, E. R., Kuge, N., & Yamamura, T. (2001). Affording realistic stopping behavior: A cardinal challenge for driving simulators. *Proceedings of the 1st human-centered transportation simulation conference*, Iowa City, Iowa.

Boer, E. R., Rakauskas, M. E., Ward, N. J., & Goodrich, M. A. (2005). Steering entropy revisited. In *Proceedings of the Third International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design* (pp. 25–32). Iowa City: University of Iowa, Public Policy Center. Retrieved from http://drivingassessment.uiowa.edu/DA2005/PDF/05_Boerformat.pdf

Bohan, M., McConnell, D. S., Chaparro, A., & Thompson, S. G. (2010). The effects of visual magnification and physical movement scale on the manipulation of a tool with indirect vision. *Journal of Experimental Psychology: Applied, 16*, 33–44.

Bohan, M., Thompson, S. G., & Samuelson, P. J. (2003). Kinematic analysis of mouse cursor positioning as a function of movement scale and joint set. *Proceedings of the International Conference on Industrial Engineering–Theory, Applications and Practice* (pp. 442–447). Wichita, KS: Wichita State University.

Boschloo, H. W., Wieringa, P. A., Kuipers, J., De Winter, J. C. F., & Mulder, M. (2005). Driving behaviour in low-cost driving simulators: The influence of vibrations on braking and cornering behaviour. *Proceedings of the 24th Annual European Conference on Human Decision Making and Manual Control*, Athens, Greece.

Breur, S. W., Pool, D. M., Van Paassen, M. M., & Mulder, M. (2010). Effects of displayed error scaling in compensatory roll-axis tracking tasks. *AIAA Modeling and Simulation Technologies Conference*, Toronto, Canada.

Brookhuis, K. A., & De Waard, D. (1993). The use of psychophysiology to assess driver status. *Ergonomics, 36*, 1099–1110.

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

Brown, Y. J., Cardullo, F. M., & Sinacori, J. B. (1989). Need-based evaluation of simulator force and motion cuing devices. In *Proceedings of AIAA flight simulation technologies conference* (AIAA 89-3272-CP). Monterey, CA: American Institute of Aeronautics and Astronautics.

Brünger-Koch, M. Briest, S., & Vollrath, M. (2006). Virtual driving with different motion characteristics: braking manoeuvre analysis and validation. *Proceedings of the Driving Simulation Conference*, Paris, France, 69–78.

Buck, L. (1980). Motor performance in relation to control-display gain and target width. *Ergonomics, 23*, 579–589.

Bürki-Cohen, J., Soja, N. N., & Longridge, T. (1998). Simulator platform motion—The need revisited. *The International Journal of Aviation Psychology, 8*, 293–317.

Bürki-Cohen, J., Sparko, A. L., & Go, T. H. (2007). Training value of a fixed-base flight simulator with a dynamic seat. *AIAA Modeling and Simulation Technologies Conference and Exhibit*, Hilton Head, South Carolina, AIAA paper 2007-6564.

Butler, J. S., Campos, J. L., & Bülthoff, H. H. (2008). The robust nature of visual-vestibular combination for heading. *Perception, 37*, ECVP Abstract Supplement, 40.

Cantin, V., Lavallière, M., Simoneau, M., & Teasdale, N. (2009). Mental workload when driving in a simulator: Effects of age and driving complexity. *Accident Analysis & Prevention, 41*, 763–771.

Caro, P. W. (1979). The relationship between flight simulator motion and training requirements. *Human Factors, 21*, 493–504.

Casanova, D., Sharp, R. S., & Symonds, P. (2002). On the optimisation of the longitudinal location of the mass centre of a Formula One car for two circuits. *Proceedings of AVEC '02* (pp. 6–12). JSAE 20024496.

Casiez, G., Vogel, D., Balakrishnan, R., & Cockburn, A. (2008). The impact of control-display gain on user performance in pointing tasks. *Human Computer Interaction, 23*, 215–250.

Cauraugh, J. H., Chen, D., & Radlo, S. (1993). Effects of traditional and reversed bandwidth knowledge of results on motor learning. *Research Quarterly for Exercise and Sport, 64*, 413–417.

Chase, R. A., Cullen, J. K., Jr., Openshaw, J. W., & Sullivan, S. A. (1965). Studies on sensory feedback: III The effects of display gain on tracking performance. *The Quarterly Journal of Experimental Psychology, 17*, 193–208.

Chung, W. W. Y., Perry Jr., C. H., & Bengford, N. J. (2001). Investigation of effectiveness of the dynamic seat in a black hawk flight simulation. *AIAA Modeling and Simulation Technologies Conference and Exhibit*, Montreal, Canada, AIAA paper 2001-4064.

Cipelli, M., Schiehlen, W., & Cheli, F. (2008). Driver-in-the-loop simulations with parametric car models. *Vehicle System Dynamics, 46*, 33–48.

Clarke, D. D., Ward, P., & Truman, W. (2005). Voluntary risk taking and skill deficits in young driver accidents in the UK. *Accident Analysis & Prevention, 37*, 523–529.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd Ed.). Hillsdale, NJ: Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159.

Colombet, F., Dagdelen, M., Reymond, G., Pere, C., Merienne, F., & Kemeny, A. (2008). Motion cueing: What is the impact on the driver's behavior? *Proceedings of the Driving Simulation Conference*, Monaco, 171–181.

Conover, W. J., & Iman, R. L. (1981). Rank transformations as a bridge between parametric and nonparametric statistics. *The American Statistician, 35*, 124–133.

Coombes, S. A., Corcos, D. M., Sprute, L., Vaillancourt, D. E. (2010). Selective regions of the visuomotor system are related to gain-induced changes in force error. *Journal of Neurophysiology, 103*, 2114–2123.

Crossman, E. R. F. W. (1959). A theory of the acquisition of speed-skill. *Ergonomics, 2*, 153–166.

Crundall, D., Underwood, G., & Chapman, P. (1999). Driving experience and the functional field of view. *Perception, 28*, 1075–1087.

Cutting, J. E. (1997). How the eye measures reality and virtual reality. *Behavior Research Methods, Instruments & Computers, 29*, 27–36.

Dahl, R. E. (2008). Biological, developmental, and neurobehavioral factors relevant to adolescent driving risks. *American Journal of Preventive Medicine, 35*, s278–s284.

Davis, B. T., & Green, P. (1995). *Benefits of sound for driving simulation: An experimental evaluation* (UMTRI-95-16). The University of Michigan, Transportation Research Institute, Ann Arbor, MI.

De Craen, S. (2010). *The X factor. A longitudinal study of calibration in young novice drivers*. Doctoral dissertation. The Netherlands: TRAIL Research School. TRAIL thesis series T2010/2.

De Groot, S., Centeno Ricote, F., & De Winter, J. C. F. (2012). The effect of tire grip on learning driving skill and driving style: A driving simulator study, *Transportation Research Part F, 15*, 413–426.

De Groot, S., De Winter, J. C. F., López-García, J. M., Mulder, M., & Wieringa, P. A. (2011). The effect of concurrent bandwidth feedback on learning the lane keeping task in a driving simulator. *Human Factors, 53*, 50–62.

De Groot, S., De Winter, J.C.F., Mulder, M., Kuipers, J., Mulder, J.A., & Wieringa, P.A. (2006). The effects of route-instruction modality on driving performance in a simulator. *Paper presented at the 9th TRAIL congress*, Rotterdam.

De Groot, S., De Winter, J. C. F., Mulder, M., & Wieringa, P. A. (2007, September). Didactics in simulator-based driver training: Current state of affairs and future potential. *Proceedings of the Driving Simulation Conference North America*, Iowa City, IA. Retrieved from http://www.nads-sc.uiowa.edu/dscna/2007/papers/Section%206A%20-%20Training/deGroot.pdf

De Groot, S., De Winter, J. C. F., Mulder, M., & Wieringa, P. A. (2011a). Car racing in a simulator: Validation and assessment of brake pedal stiffness. *Presence: Teleoperators and Virtual Environments, 20*, 47–61.

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

De Groot, S., De Winter, J. C. F., Wieringa, P. A., & Mulder, M. (2009). An analysis of braking measures. *Proceedings of the Driving Simulation Conference*, Monte Carlo, Monaco, 233–243.

De Winter, J. C. F., De Groot, S., Mulder, M., & Wieringa, P. A. (2007). The fun of engineering: a motion seat in a driving simulator. *Proceedings of the Driving Simulation Conference*, Iowa City, IA, 77–85.

De Winter, J. C. F., De Groot, S., Mulder, M., Wieringa, P. A., & Dankelman, J. (2008). The search for higher fidelity in fixed-base driving simulation: six feedback systems evaluated. *Proceedings of the Driving Simulation Conference*, Monte Carlo, Monaco, 183–192.

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*, 137–153.

De Winter, J. C. F., & Dodou, D. (2010). Five-point Likert items: t test versus Mann-Whitney Wilcoxon. *Practical Assessment, Research & Evaluation, 15*. Retrieved from http://pareonline.net/getvn.asp?v=15&n=11

De Winter, J. C. F., Dodou, D., & Mulder, M. (2012). Training effectiveness of whole body flight simulator motion: A comprehensive meta-analysis. *The International Journal of Aviation Psychology, 22*, 164–183.

De Winter, J. C. F., & Happee, R. (2012). Modelling driver behaviour: a rationale for multivariate statistics. *Theoretical Issues in Ergonomics Science, 13*, 528–545.

De Winter, J. C. F., Houtenbos, M., Wieringa, P. A., Mulder, J. A., Kuipers, J., & De Groot, S. (2006). Individual characteristics affecting intersection behaviour in a driving simulator, *Proceedings of the 25th European Conference on human decision making and manual control (EAM)*, Valenciennes, France.

De Winter, J. C. F., Wieringa, P. A., Dankelman, J., Mulder, M., De Groot, S., & Kuipers, J. (2006). Driver assessment during self and forced-paced simulation-based training. In M. P. Gauriat & A. Kemeny (Eds.), *Proceedings of the Driving Simulation Conference Europe* (pp. 157–166). Paris, France: Inrets-Renault.

De Winter, J. C. F., Wieringa, P. A., Kuipers, J., & De Groot, S. (2007). *Training on a driving simulator*, unpublished internal research report. (Delft: TU Delft).

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*, 138–158.

Drivemasters (2007). website accessed May 2007, http://www.drivemasters.nl

Drummond, A. E. (1989). *An overview of novice driver performance issues. A literature review* (Report No. 9). Clayton, Australia: Monash University, Accident Research Centre.

Eccles, D. W. (2006). Thinking outside of the box: The role of environmental adaptation in the acquisition of skilled and expert performance. *Journal of Sports Sciences, 24*, 1103–1114.

Elander, J., West, R., & French, D. (1993). Behavioral correlates of individual differences in road-traffic crash risk: An examination of methods and findings. *Psychological Bulletin, 113*, 279–294.

Ellis, R. D., Cao, A., Pandya, A., Composto, A., Chacko, M., Klein, M. et al. (2004). Optimizing the surgeon-robot interface: The effect of control-display gain and zoom level on movement time. *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting*, 1713–1717.

Elvik, R. (2004). To what extent can theory account for the findings of road safety evaluation studies? *Accident Analysis & Prevention, 36*, 841–849.

Elvik, R., Christensen, P., & Amundsen, A. (2004). *Speed and road accidents. An evaluation of the power model* (TØI report no. 740). Oslo: Institute of Transport Economics.

Elvik, R., & Vaa, T. (2004). *Handbook of road safety measures*. Oxford: Elsevier Science.

Ely, J. H., Thomson, R. M., Orlansky, J. (1956). *Design of controls* (WADC TR 56-172). Wright-Patterson Air Force Base, OH: Aero Medical Laboratory.

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences, 8*, 162–169.

Evans, L. (2004). *Traffic safety*. Bloomfield Hills, MI: Science Serving Society.

Evans, L. (2006). Innate sex differences supported by untypical traffic fatalities. *Chance, 19*, 10–15.

Falkmer, T., & Gregersen, N. P. (2005). A comparison of eye movement behavior of inexperienced and experienced drivers in real traffic environments. *Optometry & Vision Science, 82*, 732–739.

Felder R.M., & Spurlin, J. (2005) *Applications, reliability and validity of the Index of Learning Styles*, North Carolina State University, North Carolina, USA.

Finn, P., & Bragg, B. W. (1986). Perception of the risk of an accident by young and older drivers. *Accident Analysis & Prevention, 18*, 289–298.

Fisher, P., Kubitzki, J., Guter, S., & Dieter, F. (2007). Virtual driving and risk taking: Do racing games increase risk-taking cognitions, affect, and behaviors? *Journal of Experimental Psychology, Applied, 13*, 22–31.

Fisher, R., & Miller, J. (2008). Differential redundancy gain in onset detection versus offset detection. *Perception & Psychophysics, 70*, 431–436.

Fisher, D. L., Pollatsek, A. P., & Pradhan, A. (2006). Can novice drivers be trained to scan for information that will reduce their likelihood of a crash? *Injury Prevention, 12*, 25–29.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47*, 381–391.

Flach, J. M., Riccio, G. E., McMillian, G. R., & Warren, R. (1986). Psychophysical methods for equating performance between alternative motion simulators. *Ergonomics, 29*, 1423–1438.

Flach, J.M., Smith, M.R.H., Stanard, T., & Dittman, S.M. (2003). Collisions: Getting them under control. In: H. Hecht, & G.J.P. Savelsbergh (Eds.) *Time-to-Contact*. Amsterdam: Elsevier.

Frese, M., Albrecht, K., Altmann, A., Lang, J., Papstein, P. V., Peyerl, R., et al. (1988). The effects of an active development of the mental model in the training process: experimental results in a word processing system. *Behaviour and Information Technology, 7*, 295–304.

Frex Japan Trading (2008). FrexGP SimConMOTION. http://www.frex.com/gp

Fuller, R. (2005). Towards a general theory of driver behaviour. *Accident Analysis & Prevention, 37*, 461–472.

Gagné, R. (1985). *The conditions of learning* (4^{th} Ed.). New York: Rinehart and Winston.

Gagné, R. (2005). *Principles of instructional design* (5^{th} Ed.). Belmont, CA: Wadsworth.

Gavelin, A., Lindquist, M., & Oldenburg, M. (2007). Modelling and simulation of seat-integrated safety belts including studies of pelvis and torso responses in frontal crashes. *International Journal of Crashworthiness, 12*, 367–379.

Gelau, C., Sirek, J., & Dahmen-Zimmer, K. (2011). Effects of time pressure on left-turn decisions of elderly drivers in a fixed-base driving simulator. *Transportation Research Part F, 14*, 76–86.

Gibbs, C. B. (1962). Controller design: Interactions of controlling limbs, time-lags and gains in positional and velocity systems. *Ergonomics, 5*, 385–402.

Godley, S. T., Triggs, T. J., & Fildes, B. N. (2002). Driving simulator validation for speed research. *Accident Analysis & Prevention, 34*, 589–600.

Goodwin, J. E., & Meeuwsen, H. J. (1995). Using bandwidth knowledge of results to alter relative frequencies during motor skill acquisition. *Research Quarterly for Exercise and Sport, 66*, 99–104.

Gordon, N. B., & Gottlieb, M. J. (1967). Effect of supplemental visual cues on rotary pursuit. *Journal of Experimental Psychology, 75*, 566–568.

Green, P. (2005). How driving simulator data quality can be improved. *Proceedings of the Driving Simulation Conference*, Orlando, FL, 210–220.

Green Dino (2008). website accessed Jan 2008 http://www.greendino.nl

Greenberg, J., Artz, B., & Cathey, L. (2003). The effect of lateral motion cues during simulated driving. *Proceedings of the Driving Simulation Conference*, Dearborn, MI.

Gregersen, N. P. (1996). Young drivers' overestimation of their own skill-an experiment on the relation between training strategy and skill. *Accident Analysis & Prevention, 28*, 243–250.

Gregor, R. J., & Wheeler, J. B. (1994). Biomechanical factors associated with shoe/pedal interfaces. Implications for injury. *Sports Medicine, 17*, 117–131.

Groeger, J.A. (2000). *Understanding driving.* Hove, UK: Psychology press Ltd.

Groeger, J. A., & Banks, A. P. (2007). Anticipating the content and circumstances of skill transfer: Unrealistic expectations of driver training and graduated licensing? *Ergonomics, 50*, 1250–1263.

Groeger, J.A., & Clegg, B. A. (2004) *Practice and instruction when learning to drive*. Road Safety Research Report 14. London: Department of the Environment Transport and the Regions.

Guadagnoli, M. A., & Lee, T. D. (2004). Challenge point: A framework for conceptualizing the effects of various practice conditions in motor learning. *Journal of Motor Behavior, 36*, 212–224.

Hammerton, M. (1962). An investigation into the optimal gain of a velocity control system. *Ergonomics, 5*, 539–543.

Hancock, P. A. (1996). Effect of control order, augmented feedback, input device and practice on tracking performance and perceived workload. *Ergonomics, 39*, 1146–1162.

Harrison, W. A. (1999). The limited potential of training for learner drivers: A view from the psychologists' lab. *Paper presented to the Australian College of Road Safety Young Driver Conference*.

Hart, S. G. (2006). NASA-Task Load Index (NASA-TLX); 20 years later. In *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting* (pp. 904–908). Santa Monica, CA: Human Factors & Ergonomics Society.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock, & N. Meshkati (Eds.). *Human Mental Workload* (pp. 139–183). Amsterdam: North Holland B.V.

Hatakka, M., Keskinen, E., Baughan, C., Goldenbeld, C., Gregersen, N. P., Groot, H., Siegrist, S., Willmes-Lenz, G., & Winkelbauer, M. (2003). *Basic driver training: New models*. Turku, Finland: University of Turku.

Hatakka, M., Keskinen, E., Gregersen, N. P., Glad, A., & Hernetkoski, K. (2002). From control of the vehicle to personal self-control; broadening the perspectives to driver education. *Transportation Research Part F, 5*, 201–215.

Heil, D. P., Engen, J., & Higginson, B. K. (2004). Influence of ski pole grip on peak upper body power output in cross-country skiers. *European Journal of Applied Physiology, 91*, 481–487.

Heintzman, R. J. (1996). *Determination of force cueing requirements for tactical combat flight training devices* (ASC-TR-97-5001). Wright Patterson Air Force Base, OH: Training Systems Product Group.

Helson, H. (1949). Design of equipment and optimal human operation. *The American Journal of Psychology, 62*, 473–497.

Hendrix, C., & Barfield, W. (1996). Presence within virtual environments as a function of visual display parameters. *Presence: Teleoperators and Virtual Environments, 5*, 274–289.

Hess, R. A. (1973). Nonadjectival rating scales in human response experiments. *Human Factors, 15*, 275–280.

Ho, C., Reed, N., & Spence, C. (2006). Assessing the effectiveness of “intuitive” vibrotactile warning signals in preventing front-to-rear-end collisions in a driving simulator. *Accident Analysis & Prevention, 38*, 988–996.

Ho, C., Reed, N., & Spence, C. (2007). Multisensory in-car warning signals for collision avoidance. *Human Factors, 49*, 1107–1114.

Hoeschen, A., Verwey, W., Bekiaris, E., Knoll, C., Widltroither, H., De Waard, D., Uneken, E., Gregersen, N.P., Falkmer, T., & Schelin, H. (2001). *Inventory of Driver Training Needs and Major Gaps in the Relevant Training Procedures*. GRDI-1999-10024: TRAINER Deliverable No. 2.1.

Hoffman, L., & McDowd, J. M. (2010). Simulator driving performance predicts accident reports five years later. *Psychology and Aging, 25*, 741–745.

Hogarth, R. M., Gibbs, B. J., McKenzie, C. R. M., & Marquis, M. A. (1991). Learning from feedback: exactingness and incentives. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 734–752.

Hone, G. N., & Morrison, S. (1997). Accelerated training: the use of faster-than-realtime speeds can reduce the time-to-skill factor in simulator training. In *Real-time imaging II* (pp. 144–154) San Jose CA: Society of Photo-Optical Instrumentation Engineers.

Hong, S. L., Brown, A. J., & Newell, K. M. (2008). Compensatory properties of visual information in the control of isometric force. *Perception & Psychophysics, 70*, 306–313.

Hurwitz, D. S., Knodler, M. A., & Dulaski, D. M. (2005). Speed perception fidelity in a driving simulator environment. *Proceedings of the Driving Simulator Conference North America*, Orlando, FL, 343–352.

Huysmans, M. A., De Looze, M. P., Hoozemans, M. J. M., Van der Beek, A. J., & Van Dieën, H. (2006). The effect of joystick handle size and gain at two levels of required precision on performance and physical load on crane operators. *Ergonomics, 49*, 1021–1035.

Ivancic, K., & Hesketh, B. (2000). Learning from errors in a driving simulation: effects on driving skill and self-confidence. *Ergonomics, 43*, 1966–1984.

Jacobs, J. W., Prince, C., Hays, R. T., & Salas, E. (1990). *A meta-analysis of the flight simulator training research* (TR-89-006). Orlando, FL: Naval Training Systems Center.

Jacobs, P. L., Olvey, S. E., Johnson, B. M., & Cohn, K. A. (2002). Physiological responses to high-speed, open-wheel racecar driving. *Medicine & Science in Sports & Exercise, 34*, 2085–2090.

Jamson, H., & Smith, P. (2003). Are you used to it yet? Braking performance and adaptation in a fixed base driving simulator. *Proceedings of the Driving Simulation Conference North America*, Dearborn, Michigan.

Jarmasz, J. (2006). *Accelerated training for command dynamic decision making: a pilot study using microworlds*. Toronto, Canada: Defence Research and Development.

Jellinek, H. D., & Card, S. K. (1990). Powermice and user performance. *Proceedings of the CHI'90 Conference on Human Factors in Computing Systems*. New York: ACM.

Jenkins, W. L., & Connor, M. B. (1949). Some design factors in making settings on a linear scale. *Journal of Applied Psychology, 33*, 395–409.

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. Amsterdam: Elsevier.

Johnsgard, T. (1994). Fitts' law with a virtual reality glove and a mouse: Effects of gain. *Proceedings of Graphics Interface*. Burnaby, Canada: Human-Computer Communications Society.

Jonah, B. A. (1986). Accident risk and risk-taking behaviour among young drivers. *Accident Analysis & Prevention, 18,* 255–271.

Jones, B. (1995). The effectiveness of skid-car training for teenage novice drivers in Oregon. *The Chronicle of the American Driver and Traffic Safety Education Association (ADTSEA Chronicle), 43*, 3–16.

Jones, L. A., & Sarter, N. B. (2008). Tactile displays: Guidance for their design and application. *Human Factors, 50,* 90–111.

Kantowitz, B. H., & Elvers, G. C. (1988). Fitts' law with an isometric controller: Effects of order of control and control-display gain. *Journal of Motor Behavior, 20*, 53–66.

Kappé, B., & Emmerik, M.L. (2005). *The use of driving simulators for initial driver training and testing.* Report No. 75151.01. Soesterberg, The Netherlands: TNO Defence, Security and Safety.

Katila, A., Keskinen, E., & Hatakka, M. (1996). Conflicting goals of skid training. *Accident Analysis & Prevention, 28,* 785–789.

Katzourakis, D. I., De Winter, J. C. F., De Groot, S., & Happee, R. (2012). Driving simulator parameterization using double-lane change steering metrics as recorded on five modern cars. *Simulation Modelling Practice and Theory, 26*, 96–112.

Keith, N., & Frese, M. (2008). Effectiveness of error management training: a meta-analysis. *Journal of Applied Psychology, 93*, 59–69.

Kemeny, A., & Panerai, F. (2003). Evaluating perception in driving simulation experiments. *Trends in Cognitive Sciences, 7*, 31–37.

Kinkade, R. G. (1963). *A differential influence of augmented feedback on learning and on performance* (Report No. AMRLTDR-63-12). Wright-Patterson Air Force Base, OH: Aerospace Medical Research Laboratories.

Klarica, A. J. (2001). Performance in motor sports. *British Journal of Sports Medicine, 35*, 290–291.

Kraft, M., Amick, M. M., Barth, J. T., French, J. M., & Lew, H. L. (2010). A review of driving simulator parameters relevant to the Operation Enduring Freedom/Operation Iraqi Freedom veteran population. *American Journal of Physical Medicine & Rehabilitation, 89*, 336–344.

Kwon, S., Choi, E., & Chung, M. K. (2011). Effect of control-to-display gain and movement direction of information spaces on the usability of navigation on small touch-screen interfaces using tap-n-drag. *International Journal of Industrial Ergonomics, 41*, 322–330.

Laapotti, S., & Keskinen, E. (2004). Has the difference in accident patterns between male and female drivers changed between 1984 and 2000? *Accident Analysis & Prevention, 36*, 577–584.

Lai, Q., & Shea, C. H. (1999). Bandwidth knowledge of results enhances generalized motor program learning. *Research Quarterly for Exercise and Sport, 70*, 79–83.

Land, M. F., & Tatler, B. W. (2001). Steering with the head: The visual strategy of a racing driver. *Current Biology, 11*, 1215–1220.

Lee, D N. (1976), A theory of visual control of braking based on information about time-to-collision. *Perception, 5*, 437–459.

Lee, H. C. (2003). The validity of driving simulator to measure on-road driving performance of older drivers. *Journal of Transport Engineering in Australia, 8*, 89–100.

Lee, H. C., & Lee, A. H. (2005). Identifying older drivers at risk of traffic violations by using a driving simulator: a 3-year longitudinal study. *American Journal of Occupational Therapy, 59*, 97–100.

Lee, H.-C., Lee, A. H., Cameron, D., & Li-Tsang, C. (2003). Using a driving simulator to identify older drivers at inflated risk of motor vehicle crashes. *Journal of Safety Research, 34*, 453–459.

Lee, J. D. (2007). Technology and teen drivers. *Journal of Safety Research, 38*, 203–213.

Lee, J. D., Hoffman, J. D., & Hayes, E. (2004). Collision warning design to mitigate driver distraction. In *Proceedings of SIGCHI 2004 Conference on Human Factors in Computing Systems* (pp. 65–72). New York, NY: Association for Computing Machinery.

Lee, T. D., & Carnahan, H. (1990). Bandwidth knowledge of results and motor learning: More than just a relative frequency effect. *Quarterly Journal of Experimental Psychology, 42A*, 777–789.

Lee, T. D., & Maraj, B. K. V. (1994). Effects of bandwidth goals and bandwidth knowledge of results on motor learning. *Research Quarterly for Exercise and Sport, 65*, 244–249.

Lee, T. D., Swinnen, S. P., & Serrien, D. J. (1994). Cognitive effort and motor learning. *Quest, 46*, 328–344.

Lewis, C. H., & Griffin, M. J. (1977). The interaction of control gain and vibration with continuous manual control performance. *Journal of Sound and Vibration, 55*, 553–562.

Lewis-Evans, B., De Waard, D., Jolij, J., & Brookhuis, K. A. (2012). What you may not see might slow you down anyway: Masked images and driving. *PloS One, 7*, 1–14.

Lewis-Evans, B., & Rothengatter, T. (2009). Task difficulty, risk, effort and comfort in a simulated driving task—Implications for Risk Allostasis Theory. *Accident Analysis & Prevention, 41*, 1053–1063.

Liebermann, D. G., Katz, L., Hughes, M. D., Bartlett, R. M., McClements, J., & Franks, I. M. (2002). Advances in the application of information technology to sport performance. *Journal of Sports Sciences, 20*, 755–769.

Lin, J. J.-W., Duh, H. B. L., Parker, D. E., Abi-Rached, H., & Furness, T. A. (2002). Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment. *Proceedings of the IEEE Virtual Reality*, 164–171.

Lin, M. L., Radwin, R. G., & Vanderheiden, G. C. (1992). Gain effects on performance using a head-controlled computer input device. *Ergonomics, 35*, 159–175.

Lintern, G. (1980). Transfer of landing skills after training with supplementary visual cues. *Human Factors, 22*, 81–88.

Lintern, G. (1991). An informational perspective on skill transfer in human-machine systems. *Human Factors, 33*, 251–266.

Liu, Y., & Salvendy, G. (2009). Effects of measurement errors on psychometric measurements in ergonomics studies: Implications for correlations, ANOVA, linear regression, factor analysis, and linear discriminant analysis. *Ergonomics, 52*, 499–511.

Locke, E. A. (1968). Effects of knowledge of results feedback in relation to standards, and goals on reaction-time performance. *The American Journal of Psychology, 81*, 566–574.

Looney, S. W., & Jones, P. W. (2003). A method for comparing two normal means using combined samples of correlated and uncorrelated data. *Statistics in Medicine, 22*, 1601–1610.

Lund, A. K., Williams, A. F., & Zador, P. (1986). High school driver education: Further evaluation of the DeKalb County study. *Accident Analysis & Prevention, 18*, 349–357.

Ma, R., & Kaber, D. B. (2006). Presence, workload and performance effects of synthetic environment design factors. *International Journal of Human-Computer Studies, 64*, 541–552.

MacDonald, W. A., & Hoffmann, E. R. (1980). Review of relationships between steering wheel reversal rate and driving task demand. *Human Factors, 22*, 733–739.

MacKenzie, I. S., & Riddersma, S. (1994). Effects of output display and control-display gain on human performance in interactive systems. *Behaviour & Information Technology, 13*, 328–337.

Mandryk, R. L., & Gutwin, C. (2008). Perceptibility and utility of sticky targets. In *GI '08: Proceedings of Graphics Interface* (pp. 65–72). Toronto: Canadian Information Processing Society.

Mania, K., Troscianko, T., Hawkes, R., & Chalmers, A. (2003). Fidelity metrics for virtual environment simulations based on human judgments of spatial memory awareness states. *Presence: Teleoperators and Virtual Environments, 12*, 296–310.

Martin, E. A. (1986). An investigation regarding the use of a dynamic seat-pan display for training and as a device for communicating roll-axis motion information. *Aviation, Space, and Environmental Medicine, 57*, 1189–1193.

Matthews, G., Dorn, L., Hoyes, T. W., Davies, D. R., Glendon, A. I., & Taylor, R. G. (1998). Driver stress and performance on a driving simulator. *Human Factors, 40*, 137–149.

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

McCartt, A. T., Mayhew, D. R., Braitman, K. A., Ferguson, S. A., & Simpson, H. M. (2009). Effects of age and experience on young driver crashes: Review of recent literature. *Traffic Injury Prevention, 10*, 209–219.

McGehee D. V., Lee, J. D., Rizzo, M., & Bateman, K. (2001). Examination of older driver steering adaptation on a high performance driving simulator, *Proceedings of the international driving symposium on human factors in driver assessment, training and vehicle design*, Aspen Colorado, USA.

McKnight, A. J. & McKnight, A. S. (2003). Young novice drivers: careless or clueless? *Accident Analysis & Prevention, 35*, 921–925.

McLean, J. R., & Hoffmann, R. (1975). Steering reversals as a measure of driver performance and steering task difficulty. *Human Factors, 17*, 248–256.

McRuer, D. T., & Jex, H. R. (1967). A review of quasi-linear pilot models. *IEEE Transactions on Human Factors in Electronics, 8*, 231–249.

Merrill, M. D. (1997). Instructional strategies that teach. *CBT solutions, Nov./Dec.*, 1–11.

Merrill, M. D. (2002a). First principles of instruction. *Educational Technology Research and Development, 50*, 43–59.

Merrill, M. D. (2002b). A Pebble-in-the-pond model for instructional design. *Performance Improvement, 41*.

Milliken W. F., & Milliken, D. L. (1995). *Race Car Vehicle Dynamics*. Warrendale, PA, USA: Society of Automotive Engineers SAE International.

Milot, M. H., Marchal-Crespo, L., Green, C. S., Cramer, S. C., & Reinkensmeyer, D. H. (2010). Comparison of error-amplification and haptic-guidance training techniques for learning of a timing-based motor task by healthy individuals. *Experimental Brain Research, 201*, 119–131.

Minoyama, O., & Tsuchida, H. (2004). Injuries in professional motor car racing driving at a racing circuit between 1996 and 2000. *British Journal of Sports Medicine, 38*, 613–616.

Mollenhauer, M. A., Romano, R. A., & Brumm, B. (2004). The evaluation of a motion base driving simulator in a CAVE at TACOM. *Proceedings of the 24th Army Science Conference*, Orlando, FL.

Mugge, W., Schuurmans, J., Schouten, A. C., & Van der Helm, F. C. T. (2009). Sensory weighting of force and position feedback in human motor control tasks. *The Journal of Neuroscience, 29*, 5476–5482.

Mulder, M., Mulder, M., Van Paassen, M. M., Kitazaki, S., Hijikata, S., & Boer, E. R. (2004). Reaction-time task during car-following with an active gas pedal. In *Proceedings of IEEE International Conference on Systems, Man and Cybernetics* (pp. 2465–2470). Danvers, MA: IEEE. doi:10.1109/ICSMC.2004.1400700

Nakayama, O., Futami, T., Nakamura, T., & Boer, E. R. (1999). Development of a steering entropy method for evaluating driver workload. *Proceedings of the International Congress and Exposition Detroit, Michigan*. (SAE 1999-01-0892). Warrendale, PA, USA: Society of Automotive Engineers SAE International. doi:10.4271/1999-01-0892

Nashner, L. M., Black, F. O., & Wall, C. (1982). Adaptation to altered support and visual conditions during stance: patients with vestibular deficits. *The Journal of Neuroscience, 2*, 536–544.

Nikooyan, A. A., & Zadpoor, A. A. (2009). Application of virtual environments to assessment of human motor learning during reaching movements. *Presence: Teleoperators and Virtual Environments, 18*, 112–124.

Nusseck, H. G., Teufel, H. J., Nieuwenhuizen, F. M., & Bülthoff, H. H. (2008). Learning system dynamics: Transfer of training in a helicopter hover simulator. *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit* (AIAA 2008-7107). Honolulu, HI: American Institute of Aeronautics and Astronautics.

Nygren, T. E. (1991). Psychometric properties of subjective workload measurement techniques: Implications for their use in the assessment of perceived mental workload. *Human Factors, 33*, 17–33.

O’Dwyer, N. J., & Neilson, P. D. (1996). Strategic and adaptive responses to changes in a sensory-motor relation. *Human Movement Science, 15*, 745–762.

Olsen, J., Page, J., & Vulovic, Z. (2008). Initial investigations into simulating an open-wheeler racing car. *Paper presented at the Simulation Conference: Simulation - Maximising Organisational Benefits* (SimTecT), Melbourne, Australia.

Paddan, G. S., & Griffin, M. J. (2002). Effect of seating on exposures to whole-body vibration in vehicles. *Journal of Sound and Vibration, 253*, 215–241.

Parng, A. K. (1988). *An automated test of Fitts' law and effects of target width and control/display gain using a digitizer tablet* (NOSC TD 1214). San Diego, CA: Naval Ocean Systems Center.

Parrish, R. V., & Steinmetz, G. G. (1983). *Evaluation of g-seat augmentation of fixed based/moving base simulation for transport landings under two visually imposed runway width conditions* (NASA Technical Paper 2135). Washington, DC: National Aeronautics and Space Administration.

Pauzié, A. (2008). A method to assess the driver mental workload: The driving activity load index (DALI). *IET Intelligent Transport Systems, 2*, 315–322.

Pauzié, A., & Pachiaudi, G. (1997). Subjective evaluation of the mental workload in the driving context. In T. Rothengatter & E. Carbonell Vaya (Eds.), *Traffic & transport psychology: Theory and application* (pp. 173–182).Pergamon.

Peterka, R. J., & Loughlin, P. J. (2004). Dynamic regulation of sensorimotor integration in human postural control. *Journal of Neurophysiology, 91*, 410–423.

Petzoldt, T., Bär, N., & Krems, J. F. (2009). Gender effects on lane change test (LCT) performance. *Proceedings of the Fifth Annual International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design* (pp. 90–96). http://drivingassessment.uiowa.edu/DA2009/013_PetzoldtBar.pdf

Pielemeier, W. J., Jeyabalan, V., Meier, R. C., & Otto, N. C. (1997, September). Just noticeable differences in vertical vibration for subjects on an automobile seat. *Paper presented at the 32nd United Kingdom Group Meeting on Human Response to Vibration held at the ISVR*, University of Southampton, Southampton, England.

Pinto, M., Cavallo, V., Ohlmann, T., Espié, S., & Rogé, J. (2004). The perception of longitudinal accelerations: What factors influence braking manoeuvres in driving simulators? *Proceedings of the Driving Simulation Conference*, Paris, France, 139–151.

Pradhan, A. J., Hammel, K. R., DeRamus, R., Pollatsek, A., Noyce, D. A., & Fisher, D. L. (2005). Using eye movements to evaluate effects of driver age on risk perception in a driving simulator. *Human Factors, 47*, 840–852.

Prodoehl, J., & Vaillancourt, D. E. (2010). Effects of visual gain on force control at the elbow and ankle. *Experimental Brain Research, 200*, 67–79.

Ranney, T. A., Harbluk, J. L., & Noy, I. (2005). Effects of voice technology on test track driving performance: Implications for driver distraction. *Human Factors, 47*, 439–454.

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

Reed C. M. (1998). Note on Information Transfer Rates in Human Communication. *Presence: Teleoperators and Virtual Environments, 7*, 509–518.

Reed, M. P., & Green, P. A. (1999). Comparison of driving performance on-road and in a low-cost simulator using a concurrent telephone dialling task. *Ergonomics, 42*, 1015–1037.

Reeve, T. G., Dornier, L. A., & Weeks, D. J. (1990). Precision of knowledge of results: Consideration of the accuracy requirements imposed by the task. *Research Quarterly for Exercise and Sport, 61*, 284–290.

Reigeluth, C. M. (Ed.) (1999). *Instructional design theories and models: Volume 2*. London: Lawrence Erlbaum Associates.

Reimer, B., Mehler, B., Coughlin, J. F., Roy, N., & Dusek, J. A. (2011). The impact of a naturalistic hands-free cellular phone task on heart rate and simulated driving performance in two age groups. *Transportation Research Part F, 14*, 13–25.

Reymond, F., Kemeny, A., Droulez, J., & Berthoz, A. (2001). Role of lateral acceleration in curve driving: Driver model and experiments on a real vehicle and a driving simulator. *Human Factors, 43*, 483–495.

Reynolds, B., & Adams, J. A. (1953). Motor performance as a function of click reinforcement. *Journal of Experimental Psychology, 45*, 315–320.

Riecke, B. E., Schulte-Pelkum, J., Caniard, F., & Bülthoff, H. H. (2005). Towards lean and elegant self-motion simulation in virtual reality. *Proceedings of IEEE Virtual Reality*, 131–138.

Rockway, M. R. (1954). *The effect of variations in control-display ratio and exponential time delay on tracking performance* (ASR-TR-54-618). Wright-Patterson Air Force Base, OH: Aero Medical Laboratory.

Rockway, M. R., & Franks, P. E. (1959). *Effects of variations in control backlash and gain on tracking performance* (WADC TR 58-553). Wright-Patterson Air Force Base, OH: Aero Medical Laboratory.

Roenker, D., Cissell, G. M., Ball, K. K., Wadley, V. G., & Edwards, J. D. (2003). Speed-of-processing and driving simulator training result in improved driving performance. *Human Factors, 45*, 218–233.

Romiszovski, A. (1999). The development of physical skills: Instruction in the psychomotor domain. In *Instructional design theories and models*: Volume 2, C. M. Reigeluth (Ed.) (London: Lawrence Erlbaum Associates), pp. 457–481.

Rubio, S., Díaz, E., Martín, J., & Puente, J. M. (2004). Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and Workload Profile Methods. *Applied Psychology: An International Review, 53*, 61–68.

Rutten, A. D. J. (1999). Evaluation of a dynamic seat and a helmet loader in a motion-based flight simulator. *AIAA modeling and simulation technologies conference and exhibit*, Portland, OR. AIAA paper 1999–4331.

Salas, E., Bowers, C. A., & Rhodenizer, L. (1998). It is not how much you have but how you use it: Toward a rational use of simulation to support aviation training. *The International Journal of Aviation Psychology, 8*, 197–208.

Salmoni, A. W., Schmidt, R. A., & Walter, C. B. (1984). Knowledge of results and motor learning: A review and critical reappraisal. *Psychological Bulletin, 95*, 355–386.

Santos, J., Merat, N., Mouta, S., Brookhuis, K., & De Waard, D. (2005). The interaction between driving and in-vehicle information systems: Comparison of results from laboratory, simulator and real-world studies. *Transportation Research Part F, 8*, 135–146.

Särkelä, H., Takatalo, J., Komulainen, J., Nyman, G., & Häkkinen, J. (2004). Attitudes to new technology and experiential dimensions of two different digital games. *Proceedings of the Third Nordic Conference on Human-Computer Interaction*, Tampere, Finland, 349–352. Retrieved from http://s3.amazonaws.com/publicationslist.org/data/jukka.hakkinen/ref-12/p349-sarkela.pdf

Schaab, J. A., Radwin, R. G., Vanderheiden, G. C., & Hansen, P. K. (1996). A comparison of two control-display gain measures for head-controlled computer input devices. *Human Factors, 38*, 390–403.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*, 207–217.

Schmidt, R. A., & Wulf, G. (1997). Continuous concurrent feedback degrades skill learning: Implications for training and simulation. *Human Factors, 39*, 509–525.

Scott, N. W., McPherson, G. C., Ramsay, C. R., & Campbell, M. K. (2002). The method of minimization for allocation to clinical trials: a review. *Controlled Clinical Trials, 23*, 662–674.

Segel, L., & Mortimer, R. (1970). *Driver braking performance as a function of pedal-force and pedal-displacement levels* (SAE paper 700364). Warrendale, PA, USA: Society of Automotive Engineers SAE International. doi:10.4271/700364.

Sharp, R. S. (2009). Limit braking of a high-performance motorcycle. *Vehicle System Dynamics, 47*, 613–625.

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

Sherwood, D. E. (1988). Effect of bandwidth knowledge of results on movement consistency. *Perceptual and Motor Skills, 66*, 535–542.

Showalter, T.W., & Parris, B.L. (1980). *The effects of motion and g-seat cues on pilot simulator performance of three piloting tasks* (NASA Technical Paper 1601). Washington, DC: National Aeronautics and Space Administration.

Siegler, I., Reymond, G., Kemeny, A., & Berthoz, A. (2001). Sensorimotor integration in a driving simulator: contributions of motion cueing in elementary driving tasks. *Proceedings of the Driving Simulation Conference*, Sophia-Antipolis, France, 21–32.

Simons-Morton, B., Lerner, N., & Singer, J. (2005). The observed effects of teenage passengers on the risky driving behavior of teenage drivers. *Accident Analysis & Prevention, 37*, 973–982.

Smith, C. (1996). *Drive to win: The essential guide to race driving*. Warrendale, PA, USA: Society of Automotive Engineers SAE International.

Smith, P. J. K., Taylor, S. J., & Withers, K. (1997). Applying bandwidth feedback scheduling to a golf shot. *Research Quarterly for Exercise and Sport, 68*, 215–221.

Soloman B. A., & Felder, R. M. (2005). *Index of Learning Styles Questionnaire*, North Carolina, USA: North Carolina State University. Available online:http://www.engr.ncsu.edu/learningstyles/ilsweb.html.

Sosnoff, J. J., Valantine, A. D., & Newell, K. M. (2006). Independence between the amount and structure of variability at low force levels. *Neuroscience Letters, 394*, 165–169.

Sosnoff, J. J., & Newell, K. M. (2006). Information processing limitations with aging in the visual scaling of isometric force. *Experimental Brain Research, 170*, 423–432.

Stefanidis, D., Korndorffer, J. R., Markley, S., Sierra, R., Heniford, B. T., et al. (2007). Closing the gap in operative performance between novices and experts: Does harder mean better for laparoscopic simulator training? *Journal of the American College of Surgeons, 205*, 307–313.

Strayer, D. L., & Drews, F. A. (2003). Simulator training improves driver efficiency: Transfer from the simulator to the real world. *Proceedings of the Second International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, Park City, UT, 190–193.

Summala H. (1996) Accident risk and driver behaviour, *Safety Science, 22*, 103–117.

Suzuki, K., & Jansson, H. (2003). An analysis of driver's steering behaviour during auditory or haptic warnings for the designing of lane departure warning system. *Review of Automotive Engineering, 24*, 65–70.

Swinnen, S. P. (1996). Information feedback for motor skill learning: A review. In H. N. Zelaznik (Ed.), *Advances in motor learning and control* (pp. 37–66). Champaign, IL: Human Kinetics.

SWOV Institute for Road Safety Research. (2010). *SWOV fact sheet: Simulators in driver training*. Retrieved from http://www.swov.nl/rapport/Factsheets/UK/FS_Simulators_in_driver_training.pdf

Taves, D. R. (1974). Minimization: A new method of assigning patients to treatment and control groups. *Clinical Pharmacology and Therapeutics, 15*, 443–453.

Taylor, H. L., Lintern, G., & Koonce, J. M. (1993). Quasi-transfer as a predictor of transfer from simulator to airplane. *The Journal of General Psychology, 120*, 257–276.

Theeuwes, J., Alferdinck, J. W. A. M., & Perel, M. (2002). Relation between glare and driving performance. *Human Factors, 44*, 95–107.

Thompson, S. G., McConnell, D. S., Slocum, J. S. & Bohan, M. (2007). Kinematic analysis of multiple constraints on a pointing task. *Human Movement Science, 26*, 11–26.

Thorndike, E. L. (1927). The law of effect. *American Journal of Psychology, 39*, 212–222.

Tipton, C. L., & Birmingham, H. P. (1959). The influence of control gain in a first-order man–machine control system. *Human Factors, 1*, 69–71.

Turpin, D. R., Welles, R. T., & Price, C. (2007). Simulator-based learning: Achieving performance improvement independent of instructors. *Proceedings of the Fourth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, Stevenson, WA, 481–487.

Twisk, D. A. M., & Stacey, C. (2007). Trends in young driver risk and countermeasures in European countries. *Journal of Safety Research, 38*, 245–257.

Uhr, M. B. F., Felix, D., Williams, B. J., & Krueger, H. (2003). Transfer of training in an advanced driving simulator: Comparison between real world environment and simulation in a manoeuvring driving task. *Proceedings of the Driving Simulation Conference North America*, Dearborn, MI. http://www.nads-sc.uiowa.edu/dscna/2003/papers/Uhr_Transfer%20of%20Training%20in%20an%20Advanced%20Driving....pdf

Vaden, E. A., & Hall, S. (2005). The effect of simulator platform motion on pilot training transfer: A meta-analysis. *The International Journal of Aviation Psychology, 15*, 375–393.

Vaillencourt, D. E., Haibach, P. S., & Newell, K. M. (2006). Visual angle is the critical variable mediating gain-related effects in manual control. *Experimental Brain Research, 173*, 742–750.

Van der Horst, A.R.A. (1990). *A time-based analysis of road user behaviour in normal and critical encounters*. Doctoral dissertation. Delft, The Netherlands: Delft University of Technology.

Van der Kooij, H., Jacobs, R., Koopman, B., & Van der Helm, F. (2001). An adaptive model of sensory integration in a dynamic environment applied to human stance control. *Biological Cybernetics, 84*, 103–115.

Van Doorn, R. R. A., Unema, P. J. A., & Hendriks, E. J. (2005). The locus of adaptation to altered gain in aimed movements. *Human Movement Science, 24*, 31–53.

Van Erp, J. B. F. (2005). Presenting directions with a vibrotactile torso display. *Ergonomics, 48*, 302–313.

Van Erp, J. B. F., & Van Veen, H. A. H. C. (2004). Vibrotactile in-vehicle navigation system. *Transportation Research Part F, 7*, 247–256.

Van Esch, M. (2001). *Hoor ik ook wat ik zie?* Doctoral dissertation. Soesterberg, The Netherlands: Usability Engineering Group TNO.

Verster, J. C., Veldhuijzen, D. S., Patat, A., Olivier, B., & Volkerts, E. R. (2006). Hypnotics and driving safety: Meta-analyses of randomized controlled trials applying the on-the-road driving test. *Current Drug Safety, 1*, 63–71.

Vlakveld, W. P. (2011). *Hazard anticipation of young novice drivers*. Doctoral dissertation. Leidschendam, The Netherlands: SWOV Institute for Road Safety Research.

Von der Heyde, M., & Riecke, B. E. (2001). *How to cheat in motion simulation – Comparing the engineering and fun ride approach to motion cueing* (Technical Report No. 089). Tübingen, Germany: Max Planck Institute for Biological cybernetics.

Walker, M. P., Brakefield, T., Morgan, A, Hobson, J. A., & Stickgold, R. (2002). Practice with sleep makes perfect: Sleepdependent motor skill learning. *Neuron, 35*, 205–211.

Walker, S. M., Dawson, B., & Ackland, T. R. (2001). Performance enhancement in rally car drivers via heat acclimation and race simulation. *Comparative Biochemistry and Physiology - Part A: Molecular & Integrative Physiology, 128*, 701–707.

Weevers I., Kuipers, J., Brugman, A. O., Zwiers, J., Van Dijk, E. M. A. G., & Nijholt, A. (2003). The virtual driving instructor – Creating awareness in a MultiAgent system, *Proceedings of the 16th conference of the Canadian Society for Computational Studies of Intelligence,* Halifax, Canada, pp. 596–602.

Weinberg, G., & Harsham, B. (2009). Developing a low-cost driving simulator for the evaluation of in-vehicle technologies. *Proceedings of the First International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Essen, Germany, 51–54.

Weinberg, G., Harsham, B., Forlines, C., & Medenica, Z. (2010). Contextual push-to-talk: Shortening voice dialogs to improve driving performance. *Proceedings of the 12th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Mobile HCI 2010), Lisbon, Portugal, 113–122.

Wells, P., Tong, S., Sexton, B., Grayson, G., & Jones, E. (2008). *Cohort II: A study of learner and new drivers Main report* (report no. 81) (Vol. 1). London: Department for Transport.

Wickens, C. D. (1986). The effects of control dynamics on performance. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and performance Vol. II* (pp. 39-1–39-60). New York: John Wiley & Sons.

Wickens C.D., J.G. Hollands (1999). *Engineering psychology and human performance*, Prentice Hall, New Jersey, USA.

Wikipedia. (2010). rFactor. Retrieved October, 2010, from http://en.wikipedia.org/wiki/RFactor

Williams, A. F. (2006). Young driver risk factors: Successful and unsuccessful approaches for dealing with them and an agenda for the future. *Injury Prevention, 12*, i4–i8.

Williams, A. C., & Briggs, G. E. (1962). On-target versus off-target information and the acquisition of tracking skill. *Journal of Experimental Psychology, 64*, 519–525.

Winstein, C. J., Pohl, P. S., & Lewthwaite, R. (1994). Effects of physical guidance and knowledge of results on motor learning: Support for the guidance hypothesis. *Research Quarterly for Exercise and Sport, 65*, 316–323.

Wu, C., Liu, Y., & Quinn-Walsh, C. M. (2008). Queuing network modeling of a real-time psychophysiological index of mental workload—P300 in eventrelated potential (ERP). *IEEE Transactions on Systems, Man, and Cybernetics–Part A: Systems and Humans, 38*, 1068–1084.

Yeh, Y.-Y., & Wickens, C. D. (1988). Dissociation of performance and subjective measures of workload. *Human Factors, 30*, 111–120.

Yilmaz, E. M., & Warren, Jr., W. H., (1995). Visual control of braking: A test of the τ-dot hypothesis. Journal of Experimental Psychology: *Human Perception and Performance, 21*, 996–1014.

Zaal, P. M. T., Pool, D. M., Mulder, M., Van Paassen, M. M., & Mulder, J. A. (2010). Multimodal pilot model identification in real flight. *Journal of Guidance, Control & Dynamics, 33*, 1527–1538.

Zadpoor, A. A., & Nikooyan, A. A. (2010). Modeling muscle activity to study the effects of footwear on the impact forces and vibrations of the human body during running. *Journal of Biomechanics, 43*, 186–193.

Zhai, S., Accot, J., & Woltjer, R. (2004). Human action laws in electronic virtual worlds: An empirical study of path steering performance in VR. *Presence: Teleoperators and Virtual Environments, 13*, 113–127.

Zijlstra, F. R. H. (1993). Efficiency in work behaviour: A design approach for modern tools. Doctoral dissertation. Delft, The Netherlands: Delft University Press.

Acknowledgements

I would like to thank a number of people for their contributions to this thesis.

Professor Peter Wieringa and simulator developer Jorrit Kuipers, director of Green Dino BV, have initiated the Virtual Assistant project. The objective of the Virtual Assistant project was to create a learner adaptable and automated driver training system and this project has made this thesis possible.

Peter has been actively involved with the project as my promoter, has reviewed papers and helped to generate new ideas. I want to thank him for providing an open and pleasurable research environment. My second promotor, Professor Bob Mulder, has been involved in the opening stages of the project, but retired before this thesis was finished. Professor Max Mulder has taken over his duties, and he has done so with great energy. Max has helped with multiple publications and with the integration of all the papers into this thesis.

I also want to thank Jorrit for his innovative and positive attitude, for all the discussions he contributed to, and for offering the services of his company.

The Virtual Assistant project was funded by the Dutch Ministry of Economic Affairs, under its Innovation Program on Man-Machine Interaction, IOP MMI. The chairman the IOP MMI advisory committee, René Collier, also chaired the coaching committee of the Virtual Assistant project. We had meetings with the coaching committee every 6 months and I want to thank the complete committee for their efforts and continued support: Jenny Dankelman, Jorrit Kuipers, Ad van Lier, Bob Mulder, Max Mulder, Mark Neerincx, Lucas Noldus, René van Paassen, Jelke van der Pal, and Willem Vlakveld.

Major credit for his contribution to this thesis is directed to my friend and colleague Joost de Winter. Without him I would probably not have started, or finished this thesis, and it would certainly not have been as good. Joost started as my colleague and is now my copromotor.

Of my colleagues at the BioMechanical Engineering department, I want to thank Dimitra Dodou for many hours of reviewing, for the good discussions we had, and for the input she has given.

It has been a pleasure working with Riender Happee, Peter van Leeuwen, Mehdi Saffarian, and Diomidis Katzourakis. The secretary of the BioMechanical Engineering Department, Dineke Heersma, and her colleagues Anouk de Goede-Oosterhoff, Diones Supriana, and Sabrina Ramos Rodrigues have always provided support. Thank you!

Because a driving simulator is a technical instrument, it’s not surprising that many people have supplied technical support. First of all, I want to thank the staff of Green Dino BV, and especially Gijs Harmsen and Joris van de Vrande. Harm Boschloo, who started as my predecessor on the Virtual Assistant project, has contributed regularly with high quality programming support. Erik van den Berg, Eric Haardt, Thijs Papenhuizen, and Paul van der Ploeg have helped me with designing and building simulator parts and systems. My father John and brother Rolf have helped me to build the frame of the racing simulator, together with Jasper Roelen. Jan van Frankenhuyzen of the Biorobotics Lab has helped me out with miscellaneous issues on various occasions.

As a PhD researcher I had the pleasure to work with three master students, of which two contributed to the work in this thesis. Coincidentally, they were both from Spain: José Manuel López García and Fernando Centeno Ricote. Both of them did a great job.

A large number of undergraduate bachelor students have contributed to this thesis, and I want thank all of them. First of all, the eight groups who helped with the eight cueing systems of Chapter 2. Harm Boschloo and the undergraduate group who conducted the experiment of Chapter 3. Pieter van der Hammen, Cyril Kooijmans, Renee Naaktgeboren, and Bastiaan Petermeijer for their contribution to Chapter 7. Arnold de Jonge, Bart Vandewalle, and Frank van der Hulst for their efforts in setting up and conducting Experiment 1 of Chapter 8. Stefan Burger, Marcel de Graaf, Patrick Overkamp, Dion Roest, Bram Harmsen, Ard Kot, Samuel Maljaars, and Christian van Westebrugge for their contributions in programming and acting as experimenters in the tests and pilot tests of Chapter 9.

I would like to express my thanks to the Delft University Formula Student team for letting us gather valuable data of their brake systems. I also thank the Motopark Academy, Equipe Verschuur, and Carlin racing teams for their cooperation and support.

Finally, I want to thank my family and friends for their continued support, especially my parents, my wife Helen, and our children Chris and Sophie.

Curriculum vitae

Born in Haarlem, The Netherlands, 12 July 1979.

Education
1997–2006. MSc AeroSpace Engineering at Delft University of Technology. Graduated at the Design Integration and Operations of Aircraft and Rotorcraft department (DAR).

2006–2013. PhD study at Delft University of Technology, Faculty of Mechanical, Maritime and Materials Engineering, Department of BioMechanical Engineering.

Work experience
1998–2003. Racecar driver for various teams. Double national Formula Ford champion. One year of Formula 3 experience.

2003–current. Racecar engineer for various teams. Started in Formula Renault 2.0; currently working in British and European Formula 3 for Carlin Capsicum motorsport.

2012–current. Partner of SimDelft driver training; offering simulator-based training for racecar drivers.

List of Publications

Journal articles

De Groot, S., Centeno Ricote, F., & De Winter, J. C. F. (2012). The effect of tire grip on learning driving skill and driving style: A driving simulator study, *Transportation Research Part F, 15*, 413–426.

De Groot, S., De Winter, J. C. F., López-García, J. M., Mulder, M., & Wieringa, P. A. (2011). The effect of concurrent bandwidth feedback on learning the lane keeping task in a driving simulator. *Human Factors, 53*, 50–62.

De Groot, S., De Winter, J. C. F., Mulder, M., & Wieringa, P. A. (2011a). Car racing in a simulator: Validation and assessment of brake pedal stiffness. *Presence: Teleoperators and Virtual Environments, 20*, 47–61.

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

De Winter, J. C. F., & De Groot, S. (2012). The effects of control-display gain on performance of race car drivers in an isometric braking task. *Journal of sports sciences, 30*, 1747–1756.

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*, 137–153.

Katzourakis, D., De Winter, J. C. F., De Groot, S., & Happee, R. (2012). Driving simulator parameterization using double-lane change steering metrics as recorded on five modern cars. *Simulation Modelling Practice and Theory, 26*, 96–112.

Conference proceedings

Al Zaher, R., De Groot, S., Polinder, H., & Wieringa, P. A. (2010). Comparison of an axial flux and a radial flux permanent magnet motor for solar race cars. *Proceedings of the XIX International Conference on Electrical Machines (ICEM)*, Rome, Italy, 1–6.

De Groot, S., & De Winter, J. C. F. (2011). On the way to pole position: the effect of tire grip on learning to drive a racecar. *Proceedings IEEE International Conference on Systems Man and Cybernetics*, Anchorage, AL, 133–138.

De Groot, S., De Winter, J. C. F., Mulder, M., Kuipers, J., Mulder, J. A., & Wieringa, P. A. (2006). The effects of route-instruction modality on driving performance in a simulator. *Paper presented at the 9th TRAIL congress*, Rotterdam.

De Groot, S., De Winter, J. C. F., Mulder, M., & Wieringa, P. A. (2007). Didactics in simulator-based driver training: Current state of affairs and future potential. *Proceedings of the Driving Simulation Conference North America*, Iowa City, IA. Retrieved from http://www.nads-sc.uiowa.edu/dscna/2007/papers/Section%206A%20-%20Training/deGroot.pdf

De Groot, S., De Winter, J. C. F., Mulder, M., & Wieringa, P. A. (2009). The effect of brake pedal stiffness on race car driving performance. *Proceedings of the European Annual Conference on Human Decision-Making and Manual Control,* Reims, France: Université de Reims.

De Groot, S., De Winter, J. C. F., Wieringa, P. A., & Mulder, M. (2009). An analysis of braking measures. *Proceedings of the Driving Simulation Conference Europe*, Monte Carlo, Monaco, 233–243.

De Groot, S., & Pavel, M. D. (2006). New insights into pilot behaviour during hazardous rotorcraft pilot induced oscillations. *Proceedings of the American Helicopter Society annual meeting, 62*, 1130–1141.

De Winter, J. C. F., De Groot, S., Dankelman, J., Wieringa, P. A., Van Paassen, M. M., & Mulder, M. (2008). Advancing simulation-based driver training: lessons learned and future perspectives. *Proceedings of the 10th international conference on human computer interaction with mobile devices and services*, Amsterdam, The Netherlands, 459–464.

De Winter, J. C. F., De Groot, S., Mulder, M., & Wieringa, P. A. (2007). The fun of engineering: a motion seat in a driving simulator. *Proceedings of the Driving Simulation Conference North America*, Iowa City, IA, 77–85.

De Winter, J. C. F., De Groot, S., Mulder, M., Wieringa, P. A., & Dankelman, J. (2008). The search for higher fidelity in fixed-base driving simulation: six feedback systems evaluated. *Proceedings of the Driving Simulation Conference Europe*, Monte Carlo, Monaco, 183–192.

De Winter, J. C. F., Dodou, D., De Groot, S., Abbink, D. A., & Wieringa, P. A. (2009). Hands-on experience of manual control in a human-machine systems engineering course. *Proceedings of the 37th Annual SEFI Conference*, Rotterdam, the Netherlands.

De Winter, J. C. F., Houtenbos, M., Wieringa, P. A., Mulder, J. A., Kuipers, J., & De Groot, S. (2006). Individual characteristics affecting intersection behaviour in a driving simulator. *Proceedings of the 25th European conference on human decision making and manual control (EAM)*, Valenciennes, France.

De Winter, J. C. F., Mulder, M., De Groot, S., & Wieringa, P. A. (2006). Factor analysis for driving assessment applied to car following in a simulator. *Proceedings of 9th International TRAIL Congress 2006*, Rotterdam, the Netherlands.

De Winter, J. C. F., Spek, A. C. E., De Groot, S., & Wieringa, P. A. (2009). Left turn gap acceptance in a simulator: Driving skill or driving style? *Proceedings of the Driving Simulation Conference Europe*, Monte Carlo, Monaco.

De Winter, J. C. F., Wieringa, P. A., Dankelman, J., Mulder, M., De Groot, S., & Kuipers, J. (2006). Driver assessment during self and forced-paced simulation-based training. *Proceedings of the Driving Simulation Conference Europe,* Paris, France, 157–166.

De Winter, J. C. F., Wieringa, P. A., Dankelman, J., Mulder, M., Van Paassen, M. M., & De Groot, S. (2007). Driving simulator fidelity and training effectiveness. *Proceedings of the 26th European annual conference on human decision making and manual control*, Lyngby, Denmark.

Van Leeuwen, P., De Groot, S., Happee, R., & De Winter, J. C. F. (2011). Effects of concurrent continuous visual feedback on learning the lane keeping task. *Proceedings of the sixth international driving symposium on human factors in driver assessment, training and vehicle design,* Lake Tahoe, CA, 482–488.