Transportation Research Part F 15 (2012) 413–426

Contents lists available at SciVerse ScienceDirect

Transportation Research Part F

journal homepage: www.elsevier.com/locate/trf

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

S. de Groot *, F. Centeno Ricote, J.C.F. de Winter

Delft University of Technology, Faculty of Mechanical, Maritime and Materials Engineering (3mE), Department of BioMechanical Engineering (BMechE), Mekelweg 2, 2628 CD Delft, The Netherlands

ARTICLE INFO

Article history:
Received 24 January 2011
Received in revised form 19 February 2012
Accepted 22 February 2012

Keywords:
Driver training
Driving skill
Driving style
Retention performance
Driving simulator

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.

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

* Corresponding author. Tel.: +31 653384134.
E-mail addresses: s.degroot@tudelft.nl (S. de Groot), fcentenoricote@gmail.com (F. Centeno Ricote), j.c.f.dewinter@tudelft.nl (J.C.F. de Winter).

1369-8478/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.trf.2012.02.005

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.

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

Table 1
Experimental hypotheses on the relationship between lane-keeping error and speed. The expected 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

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

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.

2. Method

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.

2.2. Apparatus

The driving simulator (Fig. 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 1265 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 2100 ANSI lumens, contrast ratio 400:1, resolution 1024 $\times$ 768 pixels; side projectors NEC VT470, brightness 2000 ANSI lumens, contrast ratio 400:1, resolution 800 $\times$ 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 tires, the torque motor coupled to the steering shaft was removed. In order to produce some force feedback in the steering wheel, a spring mechanism was installed on the steering wheel axle resulting in a linear steering torque vs. steering angle relationship. Additionally, damping was applied to the steering system by wrapping a cable once around the steering shaft and attaching the two ends to opposite sides of the simulator frame; this resulted in a friction torque large enough to produce a damping ratio of about 0.5. Accordingly, the steering wheel system was modified from an active force-feedback system to a passive mass-spring-damper system. Otherwise, the high-grip group would have had much more resistance than the low-grip group, and 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 could have confounded the results. The steering sensitivity (i.e., the ratio between steering wheel angle and lateral acceleration) was calibrated to correspond to the on-center characteristics of modern cars on the road (Katzourakis, De Winter, De Groot, & Happee, in press). A number of experienced drivers tested the simulator and did not report anything unusual after driving the simulator with the passive steering system.

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 ($\leqslant$21 years vs. >21 years), (3) self-reported steering skill ($\leqslant$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.

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 as accurately as possible near the center of the right lane. All participants provided written informed consent.

Participants first performed an 8 min baseline session in the simulator to practice the auditory reaction time task (described below). Six 8 min driving sessions were then completed: four Practice sessions, an Immediate Retention session, and a Delayed Retention session on the next 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 which 22 were 90 degree curves (14 of which had a center-line radius of 20 m or less), 2 were smooth chicanes (Bezier splines), and 1 was a 180 degree curve. 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 center 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 six 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.

2.5. Dependent measures

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

2.5.1. Lane-keeping

2.5.1.1. 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 center of the right lane with zero speed and the engine switched off. Recorded data from 10 s before to 20 s after each road departure were excluded from all other measures.

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

2.5.1.3. 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, 2011).

2.5.1.4. Mean LP curves (m). The mean distance of the center of the vehicle to the center 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 center; a negative LP means that the participant drove to the right of the lane center. Mean LP left-hand curves and mean LP right-hand curves were calculated separately.

2.5.2. Speed

2.5.2.1. 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 et al., 2011; Godley, Triggs, & Fildes, 2002; Lee, Lee, Cameron, & Li-Tsang, 2003; Shechtman, Classen, Awadzi, & Mann, 2009).

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

2.5.3. Workload

2.5.3.1. Mean RT (s). The mean reaction time to 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.

2.5.3.2. 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 averaging 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).

2.5.3.3. 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".

2.5.4. Confidence

2.5.4.1. 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).

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.

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.

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	

Note: 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 Practice 4. The initial aptitude represents the covariate used in the group comparisons.

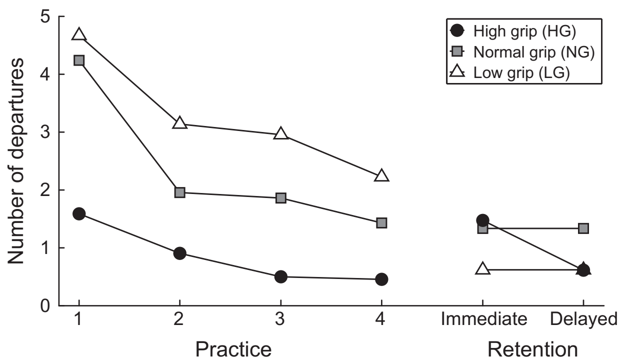

Fig. 2. Group averages of the number of road departures during the practice and retention sessions.

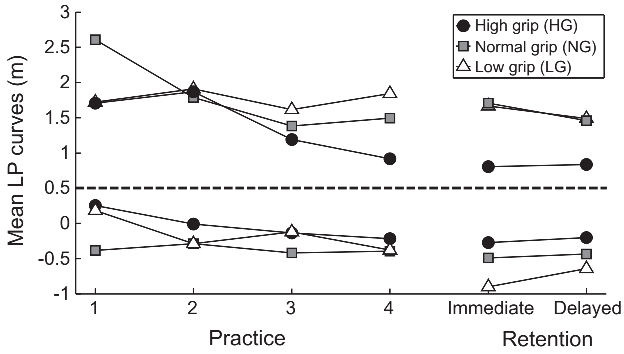

Fig. 3. Group averages of the mean lateral position (LP) in curves during the practice and retention sessions. Above the dashed line, mean LP in left curves; below the dashed line, mean LP in right curves.

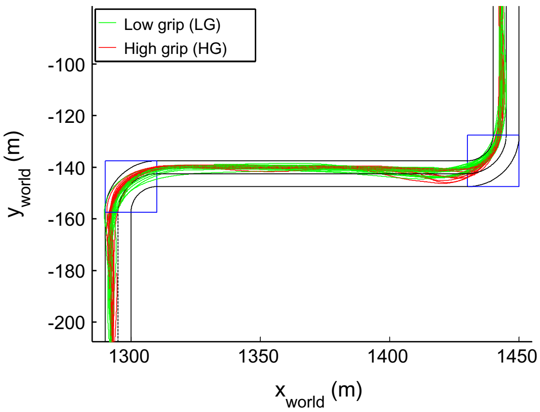

Fig. 4. Paths of the centre of the participants' car in the first two 90 degree curves of the Immediate Retention session. Participants approached from the top of this figure. Road departures were excluded from this visualization. The squares indicate the regions over which curve measures were calculated. 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.

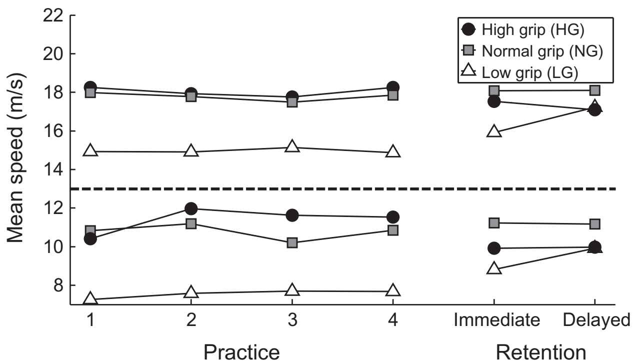

Fig. 5. Group averages of the mean speed during the practice and retention sessions. Above the dashed line, overall mean speed; below the dashed line, curve mean speed.

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 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

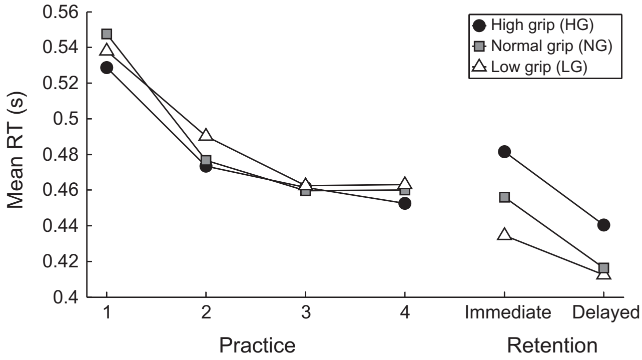

Fig. 6. Group averages of the mean reaction time (RT) during the practice and retention sessions.

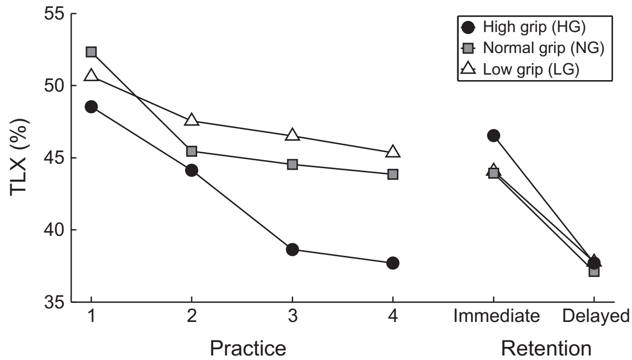

Fig. 7. Group averages of the NASA TLX score during the practice and retention sessions.

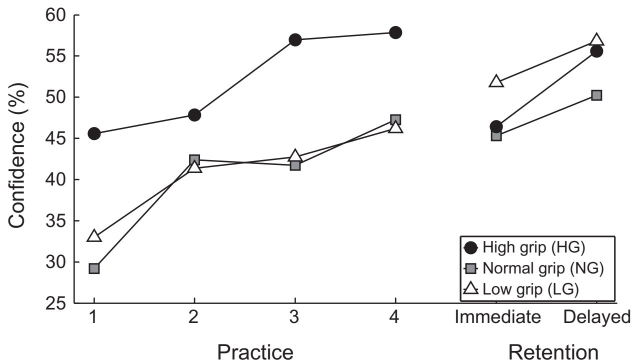

Fig. 8. Group averages of the confidence score during the practice and retention sessions.

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.

3.1. Lane-keeping

3.1.1. Number of departures

Fig. 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

Table 3
Group comparisons for the dependent measures. 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.

Measure	Session	Group comparison		
		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)

Table 4
Session differences for the three experimental groups. 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.

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

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.

3.1.2. Mean LP curves

Fig. 3 shows the Mean LP in left- and right-hand curves. The Mean LP was 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 center during the practice sessions. During Immediate Retention, HG cut the left-hand corners significantly less (i.e., drove closer to the lane center) 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. Both NG and LG improved performance in left-hand curves from Immediate to Delayed Retention. The difference between LG and HG in right-hand curves was still significant during the Delayed Retention session on the next day. Fig. 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.

3.2. Speed

3.2.1. Mean speed overall

Fig. 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. HG significantly decreased speed from Practice 4 to Immediate Retention. During Immediate Retention, LG still drove significantly slower than NG, but during Delayed Retention there was no significant difference between the three groups.

3.2.2. Mean speed curves

The trends in the group differences for the mean speed in curves (Fig. 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.

3.3. Workload

3.3.1. Mean RT

Fig. 6 shows the results for Mean RT. During practice, the Mean RT of the three groups was similar. 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.

3.3.2. TLX

Fig. 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.

3.4. Confidence

Fig. 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 a significant increase in confidence. All groups significantly increased confidence in Delayed Retention as compared to Immediate Retention. During Immediate and Delayed Retention no significant group differences were found.

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, in press; 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

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[a]	**Lower**
High grip (HG)	Mixed[a]	n.s.

[a] 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.

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, in press). 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.

Acknowledgements

Research by Joost de Winter is supported by the Stichting voor de Technische Wetenschappen (STW: the Dutch Technology Foundation), applied science division of the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO: Netherlands Organisation for Scientific Research) and the Technology Programme of the Ministry of Economic Affairs.

References

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

Allen, R. W., Park, G. D., Cook, M. L., & Fiorentino, D. (2007). *The effect of driving simulator fidelity on training effectiveness*. Paper presented at the driving simulation conference North America, Iowa City, IA.

Beckman, B. (1991). *The physics of racing. Part 5: Introduction to the racing line*. <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. *The American Journal of Occupational Therapy, 64*, 336–340.

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 and 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.

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

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

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

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

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

Dahl, R. E. (2008). Biological, developmental, and neurobehavioral factors relevant to adolescent driving risks. *American Journal of Preventive Medicine, 35*, S278–S284.

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., 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. (2011). Non-vestibular motion cueing in a fixed-base driving simulator: Effects on driver braking and cornering performance. *Presence: Teleoperators and Virtual Environments, 20*, 47–61.

De Winter, J. C. F., & Happee, R. (in press). Modelling driver behaviour: a rationale for multivariate statistics. *Theoretical Issues in Ergonomics Science*. doi: 0.1080/1463922X.2010.546437.

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., & Mulder, M. (in press). Training effectiveness of whole body flight simulator motion: A comprehensive meta-analysis. *The International Journal of Aviation Psychology*. doi:10.1080/10508414.2012.663247.

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.

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

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

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.

Elvik, R. (2004). To what extent can theory account for the findings of road safety evaluation studies? *Accident Analysis and 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.

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.

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

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.

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

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.

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

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

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.

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.

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. (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., 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.

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. *Proceedings of the SPIE, 3028*, 144–154. http://dx.doi.org/10.1117/12.270341.

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.

Jarmasz, J. (2006). *Accelerated training for command dynamic decision making: a pilot study using microworlds* (DRDC Toronto TR 2006-239). Toronto: Defence Research & Development Canada.

Jonah, B. A. (1986). Accident risk and risk-taking behaviour among young drivers. *Accident Analysis and 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.

Katila, A., Keskinen, E., & Hatakka, M. (1996). Conflicting goals of skid training. *Accident Analysis and Prevention, 28*, 785–789.

Katzourakis, D. I., De Winter, J. C. F., De Groot, S., & Happee, R. (in press). Driving simulator parameterization using double-lane change steering metrics as recorded on five modern cars. *Simulation Modelling Practice and Theory*.

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.

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.

Laapotti, S., & Keskinen, E. (2004). Has the difference in accident patterns between male and female drivers changed between 1984 and 2000? *Accident Analysis and Prevention, 36*, 577–584.

Lee, J. D. (2007). Technology and teen drivers. *Journal of Safety Research, 38*, 203–213.

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.

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 and Prevention, 41*, 1053–1063.

Lintern, G. (1991). An informational perspective on skill transfer in human-machine systems. *Human Factors, 33*, 251–266.

Lund, A. K., Williams, A. F., & Zador, P. (1986). High school driver education: Further evaluation of the DeKalb County study. *Accident Analysis and Prevention, 18*, 349–357.

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.
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.
McKnight, A. J., & McKnight, A. S. (2003). Young novice drivers: Careless or clueless? *Accident Analysis and Prevention, 35*, 921–925.
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.
Nusseck, H.-G., Teufel, H. J., Nieuwenhuizen, F. M., & Bülthoff, H. H. (2008). Learning system dynamics: Transfer of training in a helicopter hover simulator. In B. Karthik (Ed.), *Proceedings of the AIAA modeling and simulation technologies conference and exhibit (AIAA 2008-7107)*. Honolulu, HI: American Institute of Aeronautics and Astronautics.
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.
Petzoldt, T., Bär, N., & Krems, J. F. (2009). Gender effects on lane change test (LCT) performance. *Proceedings of the fifth international driving symposium on human factors in driver assessment, training and vehicle design* (pp. 90–96). <http://drivingassessment.uiowa.edu/DA2009/013_PetzoldtBar.pdf>.
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.
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.
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.
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.
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.
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.
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.
Simons-Morton, B., Lerner, N., & Singer, J. (2005). The observed effects of teenage passengers on the risky driving behavior of teenage drivers. *Accident Analysis and Prevention, 37*, 973–982.
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. In *Proceedings of the second international driving symposium on human factors in driver assessment, training and vehicle design, Park City, UT* (pp. 190–193).
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.
Turpin, D. R., Welles, R. T., & Price, C. (2007). Simulator-based learning: Achieving performance improvement independent of instructors. In *Proceedings of the fourth international driving symposium on human factors in driver assessment, training and vehicle design, Stevenson, WA* (pp. 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. In *Proceedings of the driving simulation conference North America, Dearborn, MI*.
Vlakveld, W.P. (2011). *Hazard anticipation of young novice drivers*. PhD Thesis, University of Groningen. SWOV dissertation series. Leidschendam, The Netherlands: SWOV Institute for Road Safety Research.
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.
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.
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 event-related potential (ERP). *IEEE Transactions on Systems, Man, and Cybernetics–Part A: Systems and Humans, 38*, 1068–1084.
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.