DRIVING SIMULATOR AS AN EVALUATION TOOL -ASSESSMENT OF THE INFLUENCE OF FIELD OF VIEW AND SECONDARY TASKS ON LANE KEEPING AND STEERING PERFORMANCE

Gwendolin Knappe
Andreas Keinath
Klaus Bengler
BMW Group Forschung und Technik
Germany
Cristina Meinecke
Friedrich Alexander Universität Erlangen-Nürnberg
Germany
Paper Number 07-0262

ABSTRACT

The development of new and sophisticated in-car systems fostered by technical innovation demands careful evaluation of these systems. Driving simulation is an important tool for this kind of evaluation. In-depth knowledge of the driving simulator as a tool as well as of measures recorded and calculated while using the simulator is needed to improve new driver information systems or similar devices during the development process.
For this reason, two experiments were conducted to investigate the sensitivity of lane keeping and steering measures. Participants were exposed to varying fields of view as well as cognitive and visual-motor secondary tasks.
The results yielded by the two experiments were quite consistent. All used measures are more sensitive to a visual-motor secondary task and the reduction of the peripheral field of view than to a cognitive secondary task. Out of the various steering measures the "High Frequency Component of Steering Wheel Angle" and the "Steering Wheel Reversal Rate" showed the best results. "Time to Line Crossing" and the "Standard Deviation of Lateral Position" were the most sensitive of the lane keeping measures. Since the level of difficulty in implementing and analyzing the examined measures differs widely these results can help to choose suitable measures in an economic manner. Analyses showed that a harmonization process is needed with regard to the various calculation methods of some of the measures.
Another topic was subjects' level of experience with the driving simulator. We found that only a short period of training was needed to be perfectly prepared for this kind of experiment. Interpretation of the results is limited to male persons between the age of 20 to 36 years.

INTRODUCTION

The number of built-in driving assistance and driver information systems increases continuously. Before such systems can be implemented they are tested thoroughly during the development process. These test runs ensure that driving comfort is increased without compromising safety aspects of the driver and of other traffic participants.
Driving simulation is an important tool to carry out such testing. The complexity of different driving simulators varies considerably (see Evans, 2004). Low fidelity static driving simulators consist of only a computer screen and a steering wheel as used for computer games. High fidelity driving simulators have their own mock up, the scene is extensively projected to a screen or high resolution monitors. However, a precise classification is difficult.
The dynamic driving simulator is the most complex and impressive variant, which simulates centrifugal and acceleration forces matching the according driving maneuver (see Huesmann, Ehmanns & Wisselmann, 2006).
Assessment of driving performance and driver distraction is realized by tracking eye movements, analyzing physiological measurements like pulse or heartbeat, and, probably most important, recording driving data.
A significant advantage of driving simulator tests over real-life driving tests is the fact that an expensive installation of vehicle dynamic sensors is not necessary (see Reed & Green, 1999). As a further advantage Reed and Green (1999) name the possibility to conduct standardized tests without endangering participants.
Measurements of lateral control have been used by numerous studies for a long time (see Zwahlen, Adams & DeBald, 1988, Pohlmann & Traenkle, 1994 or Pizza, Contardi, Mostacci, Mondini & Cirignotta, 2004). Many measures with various calculation methods have been suggested. Some of

these measures are recorded with different points of reference. In-depth knowledge of these points of reference and the calculation methods is necessary in order to facilitate a comparison of results across different studies.

With respect to efficient and economic test execution it is certainly advantageous to choose fitting candidates out of the set of existing measures, which are capable to show the influence of driving assistance systems on driving performance.

The presented survey deals with the systematic investigation of factors that influence driving performance when using the static driving simulator of the BMW Group. The identification of suitable lane keeping and steering wheel measures was another aim of this study. In this article findings regarding the sensitivity of measures from two experiments of the survey will be reported.

LANE KEEPING PERFORMANCE AND MEASURES OF LATERAL CONTROL

Lane keeping is a basic component of the driving task. It is the lowest level of Michon's hierarchical model (Michon, 1985). The motor and cognitive processes needed for lane keeping purposes are more or less automated at this level.

Measures of lateral control are used to describe the performance of lane keeping. They can be classified into lane keeping and steering wheel measures. Lane keeping measures are concerned with the position of the vehicle within the road or, more precisely, within a certain lane. The focus of steering wheel measures is the deviation of the steering wheel. Zwahlen, Adams & DeBald (1988) were able to show that lane keeping and steering wheel measures are sensitive to various types of distraction such as performing secondary tasks during driving.

On the basis of the norm DIN EN ISO 17287 (2003) and other surveys (see Roskam, Brookhuis, de Waard, Carsten et al., 2002), eight measures of lateral control where selected for the survey.

The chosen lane keeping measures were "Mean Lateral Position", "Standard Deviation of Lateral Position", "Time to Line Crossing" and "Number of Lane Exceedances". Steering wheel measures were "Standard Deviation of Steering Wheel Angle", "Number of Zero-Crossings", "Steering Wheel Reversal Rate" and "High Frequency Component of Steering Wheel Angle".

These measures will be briefly explained in the following sections and reasons for their use in the survey will be specified. The following explanations follow Knappe, Keinath and Meinecke (2006).

NUMBER OF LANE EXCEEDANCES (LANEX)

A lane exceedance is counted as soon as a specified part of the vehicle leaves the current lane unintentionally. In the literature, several varying definitions can be found. Östlund, Nilsson, Carsten, Merat et al. (2004) count a lane exceedance as soon as the outer side of a tire touches a lane marking. Liu, Schreiner and Dingus (1999) mention a less restrictive definition: they only talk about a lane exceedance if more than half the vehicle is on the adjacent lane.

Depending on the chosen test track the occurrence of lane exceedances might be a rare event, which is a disadvantage as it complicates the analysis of the measure. On the other hand, face validity is very high, because any lane exceedance poses a safety risk, which is why this measure was included in the survey.

For the current experiment, a lane exceedance was counted when the outer edge of either front tire exceeded the inner edge of the lane marking. All lane exceedances were counted and then divided by the distance driven (Equation 1).

$$LANEX = \frac{n_{lanex}}{d_{driven}} \quad (1).$$

This allows a comparison of results of different experiments.

MEAN LATERAL POSITION (MLP)

The mean lateral position is the average of all recorded distances (d) between a fix point of reference of the vehicle and the left or right lane boundary (Equation 2).

$$MLP = \frac{\sum_{i=1}^{n} d_i}{n} \quad (2).$$

This measure is therefore an indicator of general driving strategy or, in other words, the inclination of a driver to drift to either of the lane boundaries. When driving with extreme orientation towards one of the lane boundaries, the likelihood of a lane exceedance is increased. As de Waard, Steyvers and Brookhuis (2004) report, the lateral position might be dependent on speed: with rising speed, drivers tend to orientate towards the road center. When evaluating driving assistance and driver information systems, the question arises whether driving strategy changes while the driver uses these systems. However, a driving error can only be rated when extreme orientation towards a lane boundary is present.

The inclusion of this measure in this survey is owed to the fact that it supplies basic information about the driving strategy.

Standard Deviation of Lateral Position (SDLP)

This measure is defined as the standard deviation of all recorded distances between a fix point of reference and the left or right lane boundary (d) (Equation 3), where d_{avg} is the average of all recorded distances and n the number of distances recorded.

$$SDLP = \sqrt{\frac{\sum_{i=1}^{n}(d_i - d_{avg})^2}{n}} \qquad (3).$$

In contrast to the MLP measure, the SDLP measure is considered to judge driver distraction directly. Higher SDLP values can be interpreted as a higher deviation from the driver's chosen "ideal route" represented by the MLP. When the SDLP has very high values the probability of lane exceedances is increased. Therefore, the notion of defining driving errors based on the level of SDLP values seems justified (see Nirschl, Böttcher, Schlag & Weller, 2004).
Taking into account that the calculation of the SDLP measure is simple, it is no surprise that this measure is often included in surveys, as is the case with the paper at hand.

Time to Line crossing (TLC)

This measure was developed and specified by Godthelp, Milgram and Blaauw (1984). It specifies for a given point in time when the left or right front wheel of the vehicle would cross the lane boundary while maintaining the current course. As units of the TLC normally seconds are used. The smaller the TLC value gets, the more likely is a lane exceedance. When driving straight on a straight lane the TLC value is indefinite. Out of the recorded TLC values various TLC measures can be calculated. The simulator software calculates a TLC value for a given point in time in the following manner (Equation 4):

$$0 = \frac{(v - y - v^2 \cdot c)}{2} \cdot tlc^2 - v \cdot \alpha_{dir} \cdot tlc - d_{offs} \qquad (4).$$

In this quadratic equation v is the speed of the vehicle, y is the yaw rate of the vehicle, c is the curvature of the road, α_{dir} is the angle between vehicle and road direction and d_{offs} is the distance to the lane boundary. To obtain the TLC value, this quadratic equation can be solved with the determinant for quadratic equations.

Standard Deviation of Steering Wheel Angle (SDST)

The standard deviation of all recorded steering wheel angles is calculated to obtain this measure called SDST (see Liu et al., 1999). Although calculation of this measure is simple, the dependency on track curvature is high, which makes it difficult to sort out the influence of driver distraction. However, when comparing secondary task test runs with baseline driving, this drawback is eliminated. Since this method was employed in this survey, the according measure was also included in this study. Calculation of this measure is the same as with the SDLP measure, only with steering wheel angle deviations instead of distances.

Number of Zero-crossings (ZERO)

Each change of sign in the recorded steering wheel angle signal is counted in order to obtain the ZERO measure. The number of zero-crossings (n_{zero}) is divided by the distance driven (d_{driven}) to allow comparisons across experiments (Equation 5).

$$ZERO = \frac{n_{zero}}{d_{driven}} \qquad (5).$$

High values of this measure might indicate unstable driving behavior induced by driver distraction. However, this measure is highly influenced by track curvature like the SDST measure described in the previous section. Therefore, a comparison of task versus baseline driving is necessary. Comparisons across different surveys are only possible with accurate knowledge of track curvature (see Roskam et al., 2002).
Since the test track used in this survey was only moderately curved the measure was included despite its drawbacks.

Steering Wheel Reversal Rate (SRR)

As first mentioned by McLean and Hoffmann (1975), the calculation of this measure means a higher mathematical effort than the steering wheel measures described in the previous sections.
All reversals within the steering wheel angle signal that are greater than a given gap size are counted. The proportion of this absolute number of counted reversals (n_{gap}) and the time needed (t_{driven}) is called

steering wheel reversal rate (Equation 6).

$$SRR = \frac{n_{gap}}{t_{driven}} \quad (6).$$

In order to facilitate the determination of the reversals the steering wheel angle signal is filtered with a low pass filter, which eliminates noise in the signal. An extrema detection algorithm is employed to find minimum and maximum values in the signal. When the angle between two neighbouring extrema points is greater than the gap size, a reversal is counted.
Typically, gap sizes between a half and ten degrees are selected (see McDonald & Hoffmann, 1980). The smaller the chosen gap size, the finer the steering wheel correction that is captured with this measure. The optimal gap size has not been determined yet. Frequently, different gap sizes are used within a survey and the gap size that leads to the highest effect size is chosen. However, too large gap sizes pose the danger that reversals are only a rare event.
This measure was included in the survey to check whether the increased difficulty in obtaining the measure is worth the effort. Since the dependency on road curvature is rather low it is a promising candidate for the comparison of different surveys.

HIGH FREQUENCY COMPONENT OF STEERING WHEEL ANGLE (HFC)

McLean and Hoffman (1971) also proposed the measure called HFC. They found that steering wheel movements in a frequency band between 0.35 Hz and 0.6 Hz are sensitive for a secondary task load.
Calculation of this mathematical demanding measure is possible with different variants. According to Östlund et al. (2004), the steering wheel signal is filtered with a low pass filter (Butterworth 2nd order, cut off frequency 0.6 Hz) to eliminate noise. This filtered signal is called the "all-steering activity signal". The frequency band of interest is obtained by further filtering of the all-steering activity signal with a high pass filter (Butterworth 2nd order, cut off frequency 0.3 Hz). The HFC value is finally calculated as the proportion of the power of the frequency band signal (P_{band}) and the all-steering activity signal (P_{all}) (Equation 7).

$$HFC = \frac{P_{band}}{P_{all}} \quad (7).$$

This measure captures first and foremost high frequency steering wheel movements and thus gives information about an important aspect of steering behavior. Therefore, it was included in the survey.
Two experiments were conducted to provide a basis for publishing recommendations concerning appropriate measures in the context of evaluating driver assistance and information systems.

EXPERIMENTAL VARIATION OF THE AVAILABLE FIELD OF VIEW

The literature provides information about visual input needed for lane keeping. For example, Land & Horwood (1995) showed that a nearer part of the road (about 0.53 seconds away) is important with regard to the positioning of a car in the lane. A more distant part of the road (about one second away) gives necessary information concerning the curvature of the road. Speed plays a critical role concerning the necessary visual input. The faster a person drives, the more important is the more distant part of the road or lane-keeping performance deteriorates. Mourant and Rockwell (1972) as well as Summala, Nieminen and Punto (1996) showed how novice drivers use foveal vision for the lane-keeping task. After more driving practice has been acquired, drivers tend to use also peripheral vision. The question arises how much deterioration in the lane-keeping task occurs when peripheral vision is suppressed but still every part of the scenery can be perceived foveally.

METHOD

It was one aim of the first experiment to check whether a limitation of the field of view down to 5° degrees causes deterioration in the lane-keeping task although all parts of the road can be focused and the position of the car can also be checked foveally.
The second aim was to check whether all selected measures, including their different ways of calculation, indicate the expected change in lane keeping performance in a similar manner.

PARTICIPANTS

Twenty-one participants, mainly man, participated in the experiment. Participants were between 20 and 36 years old with an average age of 28.9 (SD = 3.9).
Participants either had normal vision or ametropia was corrected completely via contact lenses. It was not possible to wear glasses due to the experimental setup.
All participants were employees of the BMW Group and had no practical experience with driving

in a simulator before the experiment. Experiments were conducted during regular office hours; subjects participated on a voluntary basis.

APPARATUS

Driving Simulator - The static driving simulator of the BMW Group consists of a projection screen and a limousine mock-up including a roof without rear passenger area and without a trunk. The simulator is depicted in figure 1.

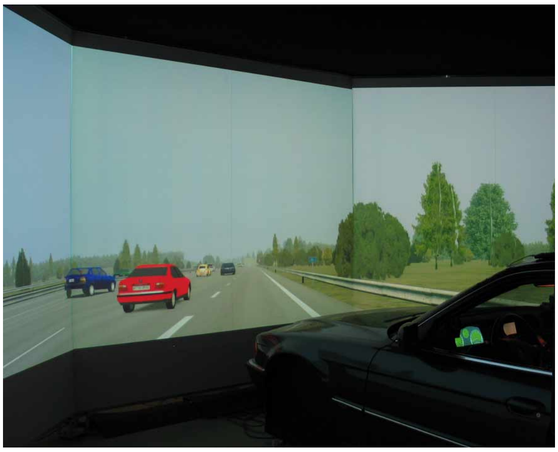

Figure 1. Static driving simulator with projection screen of the BMW Group

The projection screen consists of an angled installed screen. Three LCD projectors having a resolution of 1280x1024 pixel, project the scenery on the screen. There is a horizontal field of view of about 135° and a vertical field of view of about 38.5°. The participant is centrally seated in front of the central screen in the mock-up. The mock-up was equipped with a force-feedback steering wheel providing steering feedback depending on speed and stamped steering angle. The speedometer was fully functioning. Accelerator and brake paddle provided feedback similar to reality. By means of built-in loudspeakers driving noise was produced depending on the actual speed. Passing cars could also be determined acoustically by their simulated driving noise.

Test track - The test track used represents a 25 km long, fictitious motorway circuit, featuring three lanes in both directions. The lane width was 3.5 meters and the car width 1.89 meters. The starting point of the experimental drive was a slip road. In the experiment, participants drove the circuit anticlockwise.

The displayed scenery was to some extent slightly hilly, the maximum altitude difference being 92.3 meters with regard to the whole circuit. The scenery did not contain any hairpin bends and was just moderately curved. Different than shown in Figure 1, there were no other cars in the scenery. Other cars were not included to ensure that no additional cues for the lane-keeping task were given.

Limitation of Field of View – For methodological reasons the field of view was limited to 5° degrees via a so-called trial frame. Trial frames are glasses that can be variably adjusted and can hold glasses of different strengths. Such trial frames are used by ophthalmologists or optometrists to determine amblyopias.

For the experiment, very dark sunglasses were inserted in the trial frame. These glasses were additionally painted black on the inside. Boreholes in the center of the glasses caused a field of view of 5° degrees as depicted in figure 2.

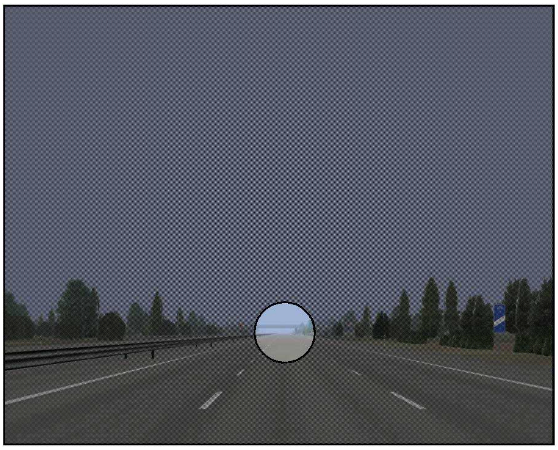

Figure 2. Limitation of the field of view

Two specially prepared “blinkers” made from robust cardboard could be mounted on the trail frame. Those “blinkers” had also a black-painted inside and prevented the lateral intrusion of light as well as the enlargement of the available field of view.

This way of limiting the field of view allowed the participants to move their head freely and to fixate any chosen part of the lane any time.

DESIGN

The size of the field of view (limited vs. standard) served as the independent variable. All lane-keeping measures described at the beginning of the paper represent the dependent measures.

Participants were assigned to two groups by random in order to eliminate order effects. Every participant took all drives under all experimental conditions. This within design excluded possible subject effects.

Procedure

The experiment was conducted at the Research and Innovation Center of the BMW Group in Munich. After a short introduction, participants got acquainted with the driving simulator during a five minutes familiarization drive. After this drive, all participants felt comfortable and perfectly prepared for the experimental drives.
Every subject participated in all drives under all experimental conditions. Half of the group started driving with just a limited field of view; the other half took the standard viewing condition first. Participants were randomly assigned to the experimental groups. The trail frame was adjusted before a trial under limited viewing conditions was started.
The instruction was given verbally via microphone and contained the following information: The participant was asked to hold a constant speed of 140 km/h while driving on a given lane. Despite the absence of real dangers, the participants were asked to drive as accurately and focused as under real driving conditions. As soon as speed exceeded or fell below the fixed speed of 140 km/h by more than 15 km/h the participant was reminded via microphone to keep a stable speed.
Due to the chosen within experimental design, the influence of single participants as source of irritation could be excluded.

Results

Data like steering wheel angle or car position within the lane were recorded with a frequency of 25 Hz.
The recorded distance between the right lane boundary and the car's center point was used to calculate MLP and SDLP. The measure SDST was calculated over all measuring points of the steering wheel angle. Out of all recorded steering wheel angle values ZERO was determined. The calculation of the measure HFC was conducted as specified by Östlund et al. (2004). Within the experiment, a LANEX was counted as soon as the outer part of a tire exceeded the lane marking of the current lane. Following Östlund et al. (2004) the gap size for the SRR measure was set at two degrees. This gap size exceeds smaller steering corrections and provides additional information with respect to other measures like the HFC measure.
Three different ways of calculating the TLC measure were used. The first one was the mean value over all local minima values (TLCmean) according to Östlund et al. (2004). For the identification of TLC minima, TLC values over 20 seconds were ignored and minima were just counted when the wave trough was broader than one second.
The second method of calculation was also suggested from Östlund et al. (2004). For this TLC measure (TLCthresh), the proportion of minima less than or equal to one second of the whole number of minima is determined. Values less than or equal to one second are considered to be especially critical as there remains almost no time for steering wheel corrections before leaving lane. Contrary to Östlund et al. (2004) minima less than or equal to two seconds were selected since minima less than or equal to one second did not occur frequently. Due to the relatively high speed of 140km/h, minima less than or equal to two seconds are regarded as critical with respect to possible lane exceedances.
Finally, the proportion of values smaller than two seconds and all values was calculated as a third method of calculation (TLCpc2).
The sensitivity of every measure concerning a limitation of the field of view was determined according to Östlund et al. (2004) by calculating Cohen's d (see Cohen, 1988). This procedure allows comparisons across different surveys. Cohen's d can be determined as soon as there is a baseline drive in addition to the experimental drive. Cohen's d is calculated as the difference of experimental drive and related baseline drive divided by their common standard deviation. According to Cohen (1988) a Cohen's d of 0.2 is considered a small effect; a Cohen's d of 0.5 or higher is considered a moderate effect. Values of 0.8 or more are considered a large effect and values greater than 1.0 describe a very large effect. The magnitude of the resulting effect size tells whether the measure in question is sensitive to a limitation of the field of view. Figure 3 depicts the results of effect size calculation.

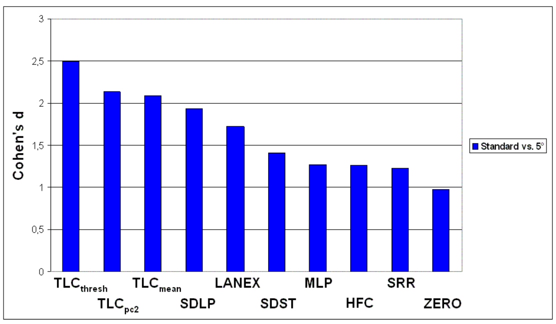

Figure 3. Effect sizes of experiment 1

All measures show large or very large effect sizes. Zero crossings show a large effect. All other measures have very large effects.

DISCUSSION

These large and very large effect sizes show that all measures in question are sensitive to a limitation of the field of view. Without peripheral vision, all measures reflect a deteriation in lane keeping performance.
However, these results are only valid for the limitation of field of view down to 5 degrees and the motorway circuit used in this experiment.
A second experiment was conducted to further examine the sensitivity of theses measures in a setting of more practical relevance.

INFLUENCE OF SECONDARY TASKS ON DRIVING PERFORMANCE

The next focus of interest was whether all measures would show similar effect sizes while the participant was carrying out secondary tasks. Based on the result, recommendations regarding measures with regard to analyzing lane-keeping performance will be derived. A further field of interest was whether a 5 minute long accommodation drive would be sufficient for driving simulator novices.

METHOD

A visual-motor and a cognitive secondary task were examined in this experiment. Regarding the visual-motor task it was of interest whether this secondary task would show a similar pattern of effect sizes as in the first experiment. Furthermore, the sensitivity for cognitive load was examined for all measures. Engström, Johansson and Östlund (2005) reported that cognitive load causes the SDLP to stabilize. When examining the results of this experiment, special attention was paid to whether this result of Engström et al. (2005) could be replicated and whether this stabilization was due to an increase in micro steering corrections as Engström suggested.

PARTICIPANTS

Twenty-nine men participated in the experiment. Age of the participants was between 22 to 36 years with a mean age being 27.2 years (SD = 3.7).
All participants were employees of the BMW Group and had no experience concerning driving simulators prior to the experiment. The experiment was conducted during regular office hours; subjects participated in the experiment voluntarily. Participants had either normal vision or brought their vision aids with them. In this case, it was also possible to wear glasses.

APPARATUS

Driving Simulator, Test track & Configuration of Traffic - This experiment used the same static driving simulator as the first experiment. The test track was also the same. Participants were supposed to drive alone in the right-hand lane with a constant speed of 120 km/h. In comparison to the first experiment, speed was reduced in order to prevent overtaxing of the participants. Other cars occupied middle and left-hand lane. Every three to five seconds those cars passed the participant's car with a speed of 130 km/h in the middle lane and 150 km/h in the left-hand lane respectively.

Secondary Tasks - One plain cognitive and one visual-motor task were chosen to judge the effects of different kinds of distraction on lane keeping performance.
The visual-motor task was taken from the ADAM project since it already proved suitable for causing visual-motor workload (see Bengler, Huesmann & Praxenthaler, 2003). Participants had to change an audiocassette while driving on the test track. This task included no cognitive aspects, as it was not necessary to keep other information such as navigation information in mind. Participants had only to perform the manual task steps and glance away from the road from time to time.
A BMW CARIN system was used for this purpose. Figure 4 shows how the system was placed in the head unit.

Figure 4. BMW CARIN system placed in the head unit

Only one button had to be pressed at the right corner of the system to open the slot and to eject the cassette.
The cognitive task required neither manual nor visual interaction with any system. Participants had to call a speech based electronic information system of the German Railway Company. They had

to find out about the arrival time of a given train at a certain station.
By using a modified head set it was ensured that no manual interaction was necessary to establish the telephone connection. Furthermore, participants were instructed to memorize arrival time and station before executing the secondary task. Thus, participants were able to keep both hands on the steering wheel during task execution.

DESIGN

The type of drive (baseline vs. visual-motor task vs. cognitive task) was the independent variable. All lane-keeping measures described at the beginning of the paper represent the dependent measures. Participants were randomly assigned to one of the six possible type pf drive orders to minimize order effects. Every participant took all drives under all experimental conditions. This within design excluded possible subject effects.

PROCEDURE

The experiment took place in the Research and Innovation Center of the BMW Group in Munich. One test run took about one hour. After a short introduction, the participants got to know the driving simulator by a five minutes familiarization drive. After this the two secondary tasks were explained. The BMW CARIN system was explained to the participant and the full and empty cassette cases used were shown. Afterwards the visual-motor task was demonstrated, then the participant was allowed to practice the task without driving: First the cassette already inserted was ejected by pressing a button in the upper right corner of the BMW CARIN system. The participant placed the cassette into the empty cassette case on the passenger seat before he removed the other cassette from its case and placed it in the cassette slot with side 2 facing up. This procedure was repeated once more before the task was completed. However, this time, the first cassette was placed with side 1 facing up into the cassette slot. The participant was instructed to begin the task on command. As soon as the participant had no further questions about the cassette task, the cognitive task was explained. Here, the participants completed the whole information dialog as an exercise. The information dialog was communicated via speech recognition to the participant.
Test runs were divided into three blocks. During each block the participant drove the same track three times, each time either performing the first, the second or no task at all.
Participants were asked to drive with a fixed speed of 120 km/h on the designated lane. They were instructed to drive as focused and carefully as when driving a real car. Furthermore, the participants were reminded that it was more important to execute the task carefully rather than quickly and that the tasks were not meant to assess their abilities. Each block was followed by a short break. As soon as speed differed by more than 15 km/h from the proposed speed, a high sound was emitted by the sound system to remind the driver to keep a stable speed. After the sound was emitted, the driver had a 10 second time frame to adjust his speed. When the speed was still not within the correct range the sound was emitted again until the speed was correct. Sound frequency held no information on the direction of speed deviation. However, the chosen frequency contrasted well to the ambient driving noise.

RESULTS

All ten measures were calculated as explained before. Only data from the second drive of every block was used for calculation. Calculation of sensitivity also follows the description given in the corresponding section of experiment 1.
Figure 5 shows an overview of the effect sizes of both the visual-motor and the cognitive task.

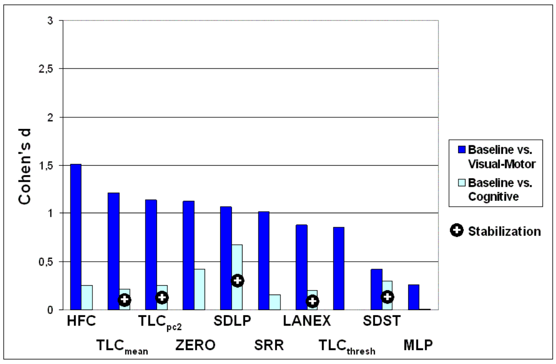

Figure 5. Effect sizes of both tasks

The visual-motor task features the first six measures showing a very large effect. The LANEX measure and the TLCthresh measure show a large effect. MLP and SDST only have a small effect.
In contrast to the visual-motor task, the cognitive task's effect sizes are overall smaller. MLP, TLCthresh and SRR have no effect. HFC, TLCmean, TLCpc2, ZERO, LANEX and SDST show small effects. SDLP has a moderate effect. A stabilization of lane keeping compared to baseline driving is found with the measures SDLP, SDST, TLCpc2, TLCmean and LANEX.
As a result of balancing the blocks, nine participants executed the baseline driving directly

after the familiarization drive. An example of the learning curve of the three drives of the block for the SDLP measure is depicted in figure 6.
Higher SDLP values represent unstable lane keeping. The line graph shows that there is no improvement for repeated baseline driving. A Friedman Test over all three baseline drives revealed no significant difference (p=0.91).

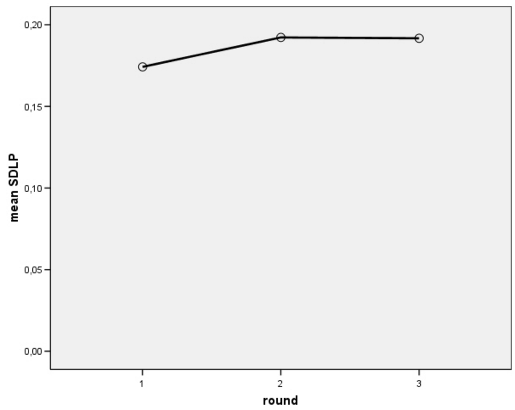

Figure 6: Learning curve for novice driving simulator drivers

DISCUSSION

The hypothesis that a visual-motor task would lead towards a measurable change in lane keeping was confirmed by eight high or very high effect sizes. Except for the MLP and the SDST, the magnitude of the effect sizes of the first and the visual-motor part of the second experiment matches remarkably well.
Since MLP does not reflect steering or lane keeping aspects but general strategy, it has a special position within the selected measures. However, a possible explanation for the lack of a significant effect of SDST might be that steering behavior while driving with unrestricted view was more or less the same regardless of task type respectively baseline driving. For the SDST to reach higher values it would be necessary to have greater steering wheel angle deviations from the average. In the second experiment, steering wheel deviations seemed to have been either equally high or equally low regardless of driving condition.
With regard to the cognitive task, all measures show no or only a small effects with the exception of the SDLP, which shows a moderate effect size. The selected measures show only a minor sensitivity for a cognitive secondary task. In other words, steering behavior while executing a cognitive task is almost the same as when performing a baseline drive. Another explanation could be that the used task was too easy.
Nonetheless, the stabilization of lane keeping found by Engström et al. (2005) can not only be seen in the SDLP measure but in LANEX, TLCmean, TLCpc2 and SDST as well.
The SDLP learning curve shows that a five-minute familiarization drive seems to be sufficient for novices to become accustomed to the static driving simulator used in this study. After this period of time no learning process can be discerned since no differences were found.

CONCLUSION

The second experiment showed that a five-minute familiarization drive is sufficient for driving simulator novices when the test track is fairly easy and no complicated maneuvers like breaking at traffic lights are required.
Additionally, all chosen lane-keeping measures proved to be sensitive to a visual-motor task as well as to a limitation of the field of view. The effect sizes are comparably high across all measures, with the exception of the measures MLP and SDST, where effect sizes were smaller for the visual-motor task.
Thus, the SDST would be an obvious candidate to omit when assessing visual-motor task influence. With respect to the other measures, a good option might be an integrative examination.
Here, additional research and/or comparison with other experiments are needed. Such a comparison might prove difficult, since calculation methods and reference points of some measures vary. In this respect, the SDLP is the least problematic measure, since the reference point is not relevant for the calculation of the standard deviation. It would facilitate matters if the calculation methods and reference points were standardized.
With regard to the cognitive task, measures proved not as sensitive as for the visual-motor task. The stabilization of lane keeping found by Engström et. al. (2005) was replicated. However, the results indicate that this stabilization is not necessarily due to increased micro steering corrections since the HFC shows only a small effect.
Due to the low sensitivity of lane keeping measures other methods such as analyses of glances, object and event detection, or measures of longitudinal control when assessing cognitive load might be preferred.
As some of these measures, for example the SDST, are more affected by road characteristics than other ones, the distribution of effect sizes across measures should be compared with results of a more curved test track.

References

Bengler, K., Huesmann, A., & Praxenthaler, M. (2003). “Investigation of Visual Demand in a Static Driving Simulator within the ADAM Project.” Proceedings of the Annual Spring Conference of the Gesellschaft für Arbeitswissenschaft e.V. (GfA) in München, 2003.

Cohen, J. (1988). “Statistical power analysis for the behavioral sciences” (2nd ed.): Lawrence Erlbaum Associates: Hillsdale, New Jersey.

de Waard, D., Steyvers, F. J. J. M., & Brookhuis, K. A. (2004). “How much visual road information is needed to drive safely and comfortably.” Safety Science, 42, 639-655.

DIN EN ISO 17287 (2003). “Straßenfahrzeuge – Ergonomische Aspekte von Fahrerinformations- und assistenzsystemen – Verfahren zur Bewertung von Gebrauchstauglichkeit beim Führen eines Kraftfahrzeugs.“

Engström, J., Johansson, E., & Östlund, J. (2005). “Effects of visual and cognitive load in real and simulated motorway driving.” Transportation Research Part F, 8, 97-120.

Evans, L. (2004). ”Traffic Safety.” Bloomfield Hills, Michigan: Science Serving Society.

Godthelp, H., Milgram, P., & Blaauw, G. J. (1984). “The development of a time-related measure to describe driving strategy.” Human Factors, 26 (3), 257-268.

Harbluk, J. L., Noy, Y. I., & Eizenmann, M. (2002). “The Impact of Cognitive Distraction on Driver Visual Behaviour and Vehicle Control.” Technical Report: TP# 13889 E, Transport Canada.

Huesmann, A., Ehmanns, D. & Wisselmann, D. (2006). “Development of ADAS by Means of Driving Simulation.” In Proceedings of the Driving Simulation Conference in Paris, 2006 (S. 131-141).

Knappe, G., Keinath, A., & Meinecke C. (2006). “Empfehlungen für die Bestimmung der Spurhaltegüte im Kontext der Fahrsimulation.“ MMI-Interaktiv Journal, 11, 3-13. http://useworld.net/servlet/handleiss?obj_id=1816&cat_id=40&vie=true (12.02.07).

Land, M., Horwood, J. (1995). “Which parts of the road guide steering?” Nature, 377(28), 339-340.

Liu, Y.-C., Schreiner, C.S., & Dingus, T.A. (1999). “Development of human factors guidelines for Advanced Traveller Information Systems (ATIS) and Commercial Vehicle Operation (CVO)”: Human Factors Evaluation of the Effectiveness of Multimodality Displays in ATIS. NHTSA FHWA-RD-96-150.

McDonald, W. A., Hoffmann, E. R. (1980). “Review of Relationships Between Steering Wheel Reversal Rate and Driving Task Demand.” Human Factors, 22(6), 733-739.

McLean, J. R., Hoffmann, E. R. (1971). “Analysis of drivers’ control movements.” Human Factors, 13, 407-418.

McLean, J. R., Hoffmann, E. R. (1975). “Steering reversals as a measure of driver performance and steering task difficulty.” Human Factors, 17, 248-256.

Michon, J.A. (1985). “A critical review of driver behavior models: what do we know, what should we do?” In L. Evans & R.C. Schwing (Ed.). Human Behavior and Traffic Safety, 485-521. New York: Plenum Press.

Mourant, R.R., Rockwell, T.H: (1972). “Strategies of visual search by novice and experienced drivers.” Human Factors, 14, 325-335.

Nirschl, G., Böttcher, S., Schlag, B., & Weller, G. (2004). “Verfahren zur Bewertung der Verkehrssicherheit von Fahrerassistenzsystemen durch objektive Erfassung von Fahrfehlerrisiken.“ In: VDI Berichte 1864. Integrierte Sicherheit und Fahrerassistenzsysteme, 397 – 420. Düsseldorf: VDI-Verlag.

Östlund, J., Nilsson, L., Carsten, O., Merat, N., Jamson, H, Jamson, S., Mouta, S., Carvalhais, J., Santos, J., Anttila, V., Sandberg, H., Luoma, J., de Waard, D., Brookhuis, K., Johansson, E., Engström, J., Victor, T., Harbluk, J., Janssen, W., & Brouwer, R. (2004) “Deliverable 2 - HMI and Safety-Related Driver Performance. Human Machine Interface And the Safety of Traffic in Europe.” Project GRD1/2000/25361 S12.319626.

Pizza, F., Contardi, S., Mostacci, B., Mondini, S., & Cirignotta, F. (2004). “A Driving Simulation Task: Correlations with Multiple Sleep Latency Test.” Brain Research Bulletin, 63(5), 423-426.

Pohlmann, S., Traenkle, U. (1994). “Orientation in Road Traffic: Age Related Differences Using an In-Vehicle Navigation System and a Conventional Map.” Accident Analysis and Prevention, 26(6), 689-702.

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(8), 1015-1037.

Roskam, A.J., Brookhuis, K.A., de Waard, D., Carsten, O.M.J., Read, L., Jamson, S., Ostlund, J., Bolling, A., Nilsson, L., Anttila, V., Hoedemaeker, M., Janssen, W.H., Harbluk, J., Johansson, E., Tevell, M., Santos, J., Fowkes, M., Engström, J. & Victor, T. (2002). “Deliverable 1 – Development of Experimental Protocol. Human Machine Interface And the Safety of Traffic in Europe (HASTE).” Project GRD1/2000/25361 S12.319626.

Summala, H., Nieminen, T., & Punto, M. (1996). “Maintaining Lane Position with Peripheral Vision during in Vehicle Tasks.” Human Factors, 38(3), 442-451.

Zwahlen, H. T., Adams, C. C. Jr., & DeBald D. P. (1988). “Safety aspects of CRT touch panel controls in automobiles.” In A. G. Gale, et al. (Ed.), Vision in Vehicles II, 335-344. Amsterdam: Elsevier Science Publishers B.V.