Head-Mounted Displays for HCI Validations While Driving

D. Reich, R. Stark

***Abstract*—**To provide reliable and valid findings when evaluating innovative in-car devices in the automotive context highly realistic driving environments are recommended. Nowadays, in-car devices are mostly evaluated due to driving simulator studies followed by real car driving experiments. Driving simulators are characterized by high internal validity, but weak regarding ecological validity. Real car driving experiments are ecologically valid, but difficult to standardize, more time-robbing and costly. One economizing suggestion is to implement more immersive driving environments when applying driving simulator studies. This paper presents research comparing non-immersive standard PC conditions with mobile and highly immersive Oculus Rift conditions while performing the Lane Change Task (LCT). Subjective data with twenty participants show advantages regarding presence and immersion experience when performing the LCT with the Oculus Rift, but affect adversely cognitive workload and simulator sickness, compared to non-immersive PC condition.

***Keywords*—**LCT, immersion, oculus rift, presence, situation awareness.

I. Introduction

USER studies are necessary to evaluate innovative in-car devices and infotainment systems in the car cockpit, e.g. touch screens, or free-hand gestures. Various methods exist for developers to validate these dual-task scenarios containing driving and interacting simultaneously, (see [1] for an overview). These methods vary significantly regarding time and costs. At this time, dual-task scenarios are mainly evaluated with driving simulator experiments. Driving simulators differ when comparing with real world driving conditions. Knappe et al. [7] summarize advantages and disadvantages of driving simulators, which should be appropriately taken into consideration when selecting a proper dual-task evaluation method:

- Potentially dangerous driving scenarios can be realized risk-free for users
- Random and singular events and traffic situations can simply be replicated
- Precisely repetition of traffic situations as often as necessary
- Simple application of more frequent unidirectional and bi-directional traffic volumes
- No restrictive factors of influence
- High effectiveness and comparability
- Less complex and more robust driving data assembly

D. R. and R. S. are with Technische Universität Berlin, Berlin, Germany 10587 (phone: 0049-030-31425419; fax: 0049-030-3930246; e-mail: diana.reich@tu-berlin.de or rainer.stark@tu-berlin.de).

Unfortunately, driving simulator studies lack regarding drivers' reality and presence experience so far. Accordingly, findings are less ecologically valid. Ecological and external validity is high for real car driving experiments, but they are only practical after the early prototyping stage has been completed. The following points also need to be taken into account when preparing a real car driving study: (a) They are difficult to control, manipulate, replicate and standardize, (b) Developers and researchers are exactingly limited when conducting real car driving experiments (e.g. physical car elements and creativity), (c) Constructing new automobiles or parts is very expensive, and (d) Real car driving experiments are more time consuming and expensive.

Our research tries to overcome the gap between these two evaluation methods by implementing more immersive parameters (e.g. 3D view, surround sound) into driving environments when applying driving simulator experiments. The more immersive a (driving) environment is, the more realistically it will be perceived by users, because of higher presence experience [9]. Here, immersion signifies the experience of submersion applied to computerized representation or simulation [10]. The five human senses (sight, sound, touch, smell, taste) should perceive the digital environment to be physically real to reach full immersion.

The new mobile device and virtual reality head mounted display Oculus Rift [4] promises to be highly immersive. Developers state that the Oculus Rift delivers a high-end virtual reality experience at an affordable price and creates a more natural and comfortable experience within virtual worlds and games. This is the reason why we wanted to examine the potential of the Oculus Rift while performing the LCT. Additionally, by using a mobile device, in this case a head mounted display, instead of a fixed display-architecture, user studies will become more flexible (e.g. regarding situations and choice of participants).

This paper discusses research comparing the influence of non-immersive (PC) conditions and highly immersive (Oculus Rift) conditions while performing the standardized LCT.

II. Related Work

The influence of different immersive parameters (visibility and sound) on human behavior was analyzed, previously [13]. Therefore, visual parameters (2D vs. 3D view) and auditory parameters (no sound vs. car sound) were used to examine different levels of immersive environments. The participants had to perform the LCT in a driving simulator. Findings show advantages for the most immersive driving environment (3D view; car sound) regarding subjective user experience ratings of the environment, and provide evidence for a more aware and realistic perception of the driving situation. Regarding situations awareness measured by SAGAT (Situation Awareness Global Assessment Technique, [2]) no statistical differences were found, probably due to realization issues. Up to that point, there was no possibility to modify the existing LCT, e.g. adding peripheral stimuli in the driving scene. Unfortunately, this is very important to measure situation awareness with the SAGAT method. Therefore, our aim was to re-build the LCT with an open source driving simulator software and add these relevant stimuli into the driving scene.

The second driving simulator study investigated the influence of immersive driving environments while interacting with a navigation system via different interaction modalities [12]. Two different immersive levels (low vs. high) and three interaction modalities (touch; spin controller; free-hand gestures) were used. Results showed significant differences in gaze data regarding low and high immersive conditions [11]. These effects were found for touch interactions relating to most glance parameters on the AOI `navigation system`. Findings showed that driving and interacting within high immersive condition leads to less visual allocation on the navigation system interface, compared to low immersive conditions. This is an indirect hint for higher situation awareness within more immersive environments. Another important finding of this study were subjective user ratings regarding users presence and immersion experiences. They were significantly higher within the high immersive driving condition (3D view; car sound).

III. Method

The experiment was realized as completely crossed 2x2 within-subjects design ($N = 20$). The first independent variable was *mode of visualization* (non-immersive vs. highly immersive). Hereby, the highly immersive mode of visualization (Oculus Rift) was characterized by stereoscopic 3D view, stereo car sound, 360° head tracking, and approximately 100° field of view (see Fig. 1 (a)). The non-immersive mode of visualization (PC) was characterized by 2D view, stereo car sound, no head tracking, and a 75° field of view (see Fig. 1 (b)). The second independent variable was *type of stimuli* (driving-relevant vs. driving-irrelevant).

The study consists of four experimental conditions and all participants had to pass the driving tasks under randomized conditions: (1) non-immersive; driving-irrelevant, (2) non-immersive; driving-relevant, (3) high immersive; driving-irrelevant and, (4) high immersive; driving-relevant. Dependent variables were (A) Situation awareness, (B) Presence and Immersive Tendencies, (C) Cognitive workload and, (D) Simulator sickness.

The German questionnaire for presence and immersive tendency in virtual realities (PIT) was used for analyzing the influence of non-immersive and highly immersive conditions on user experiences, [14]. The PIT contains three main dimensions: (1) spatial presence, (2) quality of the interface, and (3) involvement. Cognitive workload was collected by the NASA Task Load Index (NASA-TLX; [3]). The questionnaire comprises six subscales: (1) mental demand, (2) physical demand, (3) temporal demand, (4) task performance, (5) effort and (6) stress. The Simulator Sickness Questionnaire (SSQ) was used to monitor the subjects' up-to-date state [6]. Sixteen symptoms of nausea and dizziness were assessed in the form of a checklist, e.g. fatigue, headache, eyestrain.

The experiment was realized with 20 participants (*male* = 17; *female* = 3) with an average age of 28.3 years (*SD* = 3.6). Approximately half of the subjects were students ($N = 11$), the remaining nine were employed. Only one subject was left-handed. Five participants wore glasses and another two wore contact lenses. Only eight participants had experiences with head mounted displays, the remaining 12 were unexperienced.

(a) (b)

Fig. 1 User task conditions (a) Immersive with Oculus Rift (b) Non-immersive PC set-up

A. User Tasks

The user task was driving and performing the LCT, a standardized and ISO-normed tool for detecting driver distraction [5]. The task contains of a sequence of lane change maneuvers while driving with a permanent speed of 60 km/h on a three-lane road. To perform the LCT simple hardware (PC and game steering wheel) is sufficient. In this study, it was controlled via Logitech Driving Force GT force feedback wheel system which included a steering wheel and gas and brake pedal (see Fig. 1). This was connected to the PC and Oculus Rift. We rebuilt the LCT with the programmable open source driving simulator OpenDS [8]. The software is programmed entirely in Java and is based on the JMonkeyEngine framework, a video game engine. For additional details, see the OpenDS home page (http://www.opends.eu).

The situation awareness task was measured by SAGAT *(Situation Awareness Global Assessment Technique,* [2]) during each LCT condition. The participants were asked to stop four times per run at varying, unpredictable moments. The screen (driving environment) was turned off and participants were asked to answer situational questions, e.g. "What did you see at the roadside?" "What was the color of the object?" or "What is your actual lane positon?" After answering the questions the screen was turned back on and the subject completed the LCT. The answers could only be right or wrong, which determined the corresponding values of 1 or 0 as results. To indicate the level of situation awareness the values were added. For analysis the numbers of right answers were counted (0-8). The *driving-relevant* stimuli were: (1) deer, (2) baby buggy, (3) person, and (4) road sign. The *driving-irrelevant* stimuli were: (1) triangle, (2) circle, (3) square, and (4) cone.

B. Procedure

The study was performed in the VRSC (Virtual Reality Solution Center) at the Fraunhofer IPK Berlin. After the welcome by the experimenter and the registering of demographic data by means of a demographic questionnaire and briefly instruction, the current condition of the subject was determined via SSQ. Thereafter, the participants had to pass the LCT under four randomized driving conditions. In order to assess the situation awareness each run was interrupted four times for SAGAT questions. After completion of each run (condition), the subjects were asked regarding their cognitive workload. After every second immersive or non-immersive condition they had to fill out the questionnaire for presence and immersive tendency and SSQ. After receiving some sweets the participants were released.

IV. Findings

Weighted presence and immersive tendency scores were compared across conditions using paired t-tests with a significance level of .05. The t-tests showed significant differences between PC and Oculus Rift conditions, $t(16) = -2.18$, $p = .045$ for *spatial presence* ratings. Fig. 2 shows that they were highest when performing the LCT within highly immersive Oculus Rift conditions ($M = 3.19$; $SD = .46$) compared to the non-immersive standard PC conditions ($M = 2.86$; $SD = .76$). The other two dimensions, *quality of the interface* and *involvement* showed the same descriptive statistics, however they didn`t become significant.

The scores of the SAGAT data were compared across conditions using a repeated measures 2x2 ANOVA with a Greenhouse-Geisser correction. The two main factors were *mode of visualization* and *type of stimuli* (see Fig. 3).

There was found a significant main effect for *mode of visualization*, $F(1,18) = 7.23$, $p = .015$, $\eta^2 = .287$. Another significant main effect for *type of stimuli* was found, $F(1,18) = 19.00$, $p = .000$, $\eta^2 = .514$. A significant interaction between the type of visualization and stimuli was also found, $F(1,18) = 14.57$, $p = .001$, $\eta^2 = .447$. Fig. 3 shows that participants answered SAGAT questions best when performing the LCT within non-immersive PC conditions and irrelevant stimuli ($M = 7.5$; $SD = .82$).

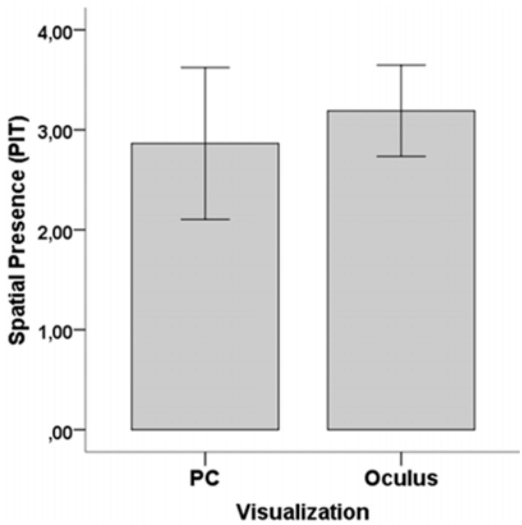
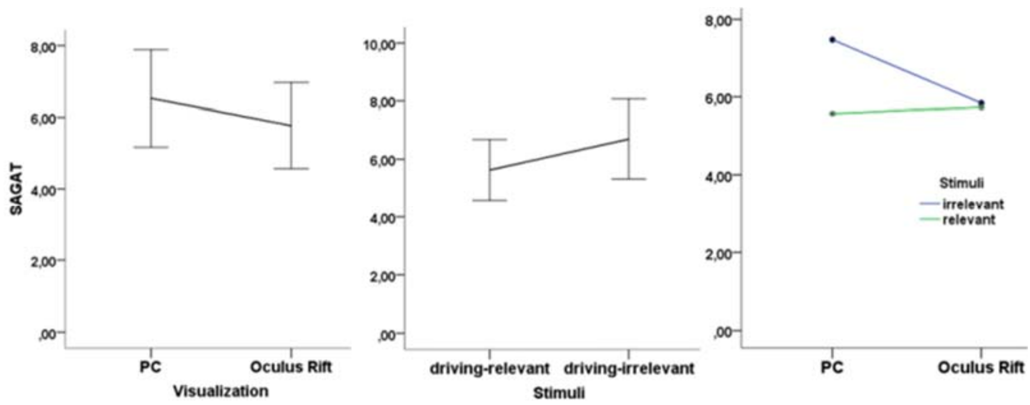
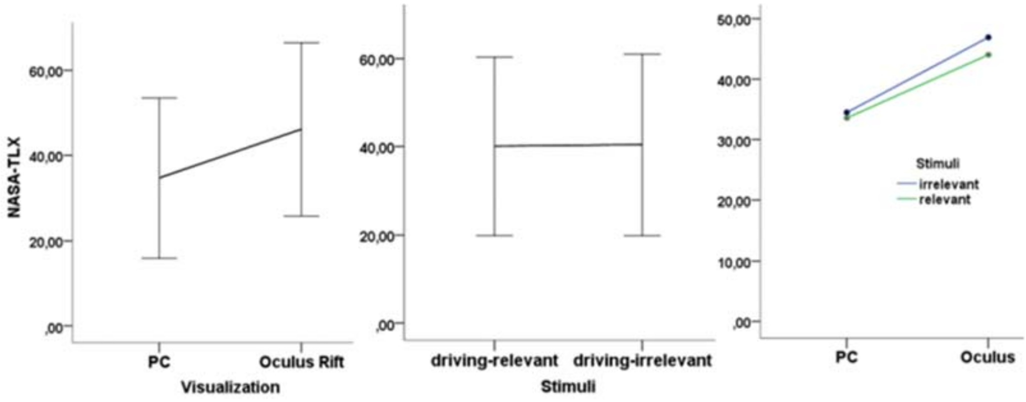


Fig. 2 Main effects for spatial presence ratings (PIT) Situation Awareness



(a) (b) (c)

Fig. 3 Main effect for mode of visualization (a), Main effect for type of stimuli (b), Interactions for situation awareness ratings (c)



(a) (b) (c)

Fig. 4 Main effect for mode of visualization (a), Main effect for type of stimuli (b), Interactions for cognitive workload ratings (c)

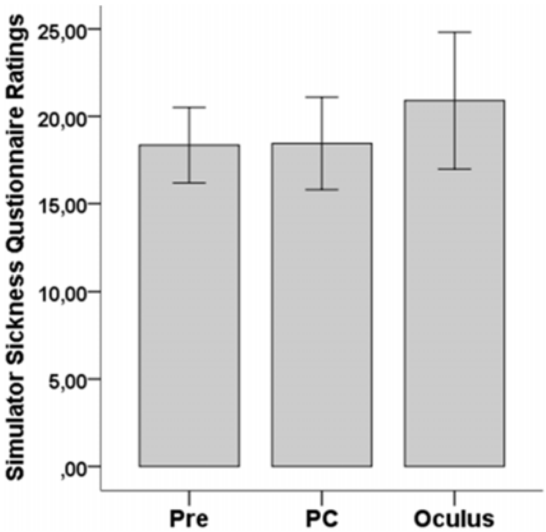


Fig. 5 Main effects for simulator sickness questionnaire ratings

B. Cognitive Workload

Weighted NASA-TLX scores were compared across conditions using a repeated measures 2x2 ANOVA with a Greenhouse-Geisser correction. The two main factors were *mode of visualization* and *type of stimuli* (see Fig. 4). There was found a significant main effect for *mode of visualization*, $F(1,17) = 18.23$, $p = .001$, $\eta^2 = .518$. Cognitive workload was lower within non-immersive PC conditions ($M = 34.48$; $SD = 19.44$), compared to highly immersive conditions ($M = 46.91$; $SD = 21.21$). There was no other significant main effect found for *type of stimuli* or an interaction between both factors.

C. Simulator Sickness

Fig. 5 shows weighted SSQ scores, which were compared across conditions using a repeated measured ANOVA, with a Greenhouse-Geisser correction. The ANOVA showed significant differences between pre, PC and Oculus Rift conditions, $F(2,38) = 4.96$, $p = < .05$, $\eta^2 = .207$. Simulator sickness symptoms were lowest when participants arrived (pre) ($M = 18.35$; $SD = 2.15$) and almost similar for non-immersive PC conditions ($M = 18.45$; $SD = 2.65$), compared to high immersive conditions ($M = 20.90$; $SD = 3.92$), where simulator sickness symptoms were significantly higher.

V. Discussion

This paper discusses research comparing the influence of non-immersive and highly immersive driving environments while performing the LCT. On the one hand, two different modes of visualization (PC; Oculus Rift) and on the other hand two different types of presented stimuli (relevant; irrelevant) were analyzed and compared.

Findings showed that performing the LCT within these different immersive experimental conditions leads to significant differences regarding subjective and objective data. Subjective user ratings of spatial presence experience were significantly higher when performing the LCT within highly immersive (Oculus Rift) conditions, characterized by stereoscopic view and head tracking, compared to non-immersive standard PC conditions with 2D front view. Due to the possibility of these mobile devices on users head they are able to completely forget the "real" environment and focus primarily on the virtual presented scene. This is a great opportunity for user testing. Nevertheless, cognitive workload and simulator sickness were significant lower when performing the LCT within non-immersive experimental standard PC set-ups. In Fact, some participants reported awkwardness when wearing a closed display on their head, which does not allow them to see their own hands on the steering wheel while driving a (virtual) car. Thus, this leads us to the conclusion, that they felt more aware of the driving situation and anticipated risks within more immersive environments. Although, it might be beneficial when performing the driving task within an immersive environment which enables drivers to see their own hands on the wheel, e.g. a Cave Automatic Virtual Environment (CAVE) or Powerwall.

Results showed significant differences regarding situation awareness when analyzing SAGAT data. Unexpectedly, participants answered more situational questions correctly within the non-immersive conditions and with irrelevant stimuli - probably related to configuration issues. One possible explanation: We tried to scale every stimuli the same size, but we should have better chosen the volume of the figures instead. Thus, the relevant stimuli “person” had much lower geometry volume on the screen compared to the same size irrelevant stimuli “circle”. Therefore, it was probably much easier for participants to recognize this one than the slimmer person. Another explanation could be that the blue circle was more salient as the relevant stimuli were (e.g. beige-colored person).

VI. Outlook

Future work will examine the same experimental set-up within another high immersive driving environment, where drivers are able to see their own body and car actuators. This high immersive driving environment consists of a 360° 3D view CAVE and a dynamic driving simulator (see http://digitalcubetestcenter.de/ for further details). Therefore, it is possible to analyze other immersive influencing factors (e.g. haptic or tactile feedback), which may also affect presence experience positively, too.

The last step will be a real car driving experiment, which will provide data for comparison. That investigation will encourage to further check the validity of the evaluation methodology within immersive environments.

References

[1] J. Breuer, K. Bengler., C. Heinrich, and W. Reichelt, “Development of advanced driver attention metrics (ADAM).” In H. Strasser, K. Kluth, H. Rausch, & H. Bubb (Eds.). Quality of Work and Products in Enterprises of the Future. Proceedings of the 50th – Anniversary Conference of the GfA and the XVII Annual ISOES Conference in Munich, 2003, pp. 37-40. Stuttgart: Ergonomia.

[2] M. R. Endsley, “Measurement of situation awareness in dynamic systems” Human Factors, 37(1), 1995a, 65-84.

[3] S. Hart and L. Staveland, “Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research”, In P. Hancock & N. Meshkati (Eds.), Human mental workload, 1988. pp. 139-183. Amsterdam: North Holland.

[4] https://www.oculus.com/rift/ (last update: 16.04.2015)

[5] ISO, *Road vehicles – Ergonomic aspects of transport information and control systems - Simulated lane change test to assess driver distraction*, ISO/TC 22/SC 13 N WG8 N579, DIS 26022., 2008.

[6] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, “Simulator Sickness Questionnaire (SSQ): An Enhanced Method for Quantifying Simulator Sickness.” International Journal of Aviation Psychology, 3 (3), 1993, 203-220.

[7] G. Knappe, A. Keinath, and C. Meinecke. „Empfehlungen für die Bestimmung der Spurhaltegüte im Kontext der Fahrsimulation.“ In M. R.K. Baumann, S. Leuchter, & L. Urbas (Hrsg.). MMI-Interaktiv, Aufmerksamkeit und Situation Awareness im Fahrzeug, 11, 2006, ISSN 1439-7854, 3-13.

[8] R. Math, A. Mahr, M. Moniri, and C. Müller, “OpenDS: A New Open-source Driving Simulator for Research”, GMM-Fachbericht-AmE, 2013.

[9] R. Moreno, and R. E. Mayer, “Learning Science in Virtual Reality Multimedia Environments: Role of Methods and Media” Journal of Educational Psychology, 94(3), 2002, 598-610.

[10] J. Nechvatal, *Immersive Ideals/Critical Distances*, LAP Lambert Academic Publishing, 2009, pp. 48-60.

[11] D. Reich, and R. Stark, “`To See or not to See?` How Do Eye Movements Change Within Immersive Driving Environments”, In Adjunct Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI,14), September 17–19, 2014, Seattle, WA, USA. http://dx.doi.org/10.1145/2667239.2667268

[12] D. Reich, and R. Stark, “The Influence of Immersive Driving Environments on Human-Cockpit Evaluations”, Hawai’i International Conference on System Sciences, Kauai January 5-8, 2015.

[13] D. Reich, H. J. Wittkugel, and R. Stark. „How Do Immersive Driving Environments Influence User Performances and Experiences?”, In T. Ahram, W. Karwowski and T. Marek. In Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics AHFE, 2014.

[14] R. Scheuchenpflug, Measuring presence in virtual environments. In M.J. Smith, G. Salvendy, & M. R. Kasdorf (Eds.). HCI International, 2001, pp. 56-58. New Orleans: HCI International.