Empirical Investigation of Transferring Cockpit Interactions from Virtual to Real-Life Environments

Diana Reich and Elisabeth Dittrich

Research Training Group prometei, Berlin, Germany
{diana.reich,elisabeth.dittrich}@zmms.tu-berlin.de

Abstract. Human-cockpit interaction is an innovative and promising field of automotive research. Indeed, automakers need to ensure safety and user satisfaction for their cockpit development concepts, if driving and interacting occurs simultaneously. One suggested approach is to evaluate simple cockpit prototypes within virtual test environments. Hybrid prototyping allows a more realistic experience with the prototype in early stages of development. With our research study we focused on important basic parameters within hybrid test environments (e.g. shutter glasses and virtual projected car model) and evaluated their potential of influence. There are no hints to assume that shutter glasses influence user behaviour. Interestingly, we found significant faster task completion times within a virtual projected car model, which indicate that immersive environments increase user performances. In summary, we can suggest hybrid prototyping within immersive test environments for evaluating human-cockpit interactions.

Keywords: human-cockpit interaction, virtual environment, hybrid prototyping.

1 Introduction

There are different research and development concepts in the area of car cockpits (Ablaßmeier, 2009). The „Cockpit of the Future" is a hot topic and on every automakers lips. Regarding to different automotive literature research, there are three big and influential development streams in the field of innovative human-cockpit systems, beside voice entries: touch (e.g. MyFord Touch), spin control (e.g. BMWs iDrive) and gestures (e.g. Daimlers Dice). Accordingly, different cockpit concepts and interfaces require specific ways of human-cockpit interaction (Broy, 2007). Because of different interaction techniques, divergent cognitive processes and parameters get relevant, e.g. distribution of attention and situation awareness (Jahn et al., 2005). But how can industry, developers and researchers define the gold standard for their systems regarding human abilities?

Not all of these innovative and for developers seemingly usable and intuitive cockpit concepts are in fact user-friendly: while driving and interacting with the cockpit

R. Shumaker (Ed.): VAMR/HCII 2013, Part II, LNCS 8022, pp. 301–309, 2013.

simultaneously difficulties can arise easily. Different studies reported negative ratings and user experiences concerning usability, user satisfaction and distribution of attention while driving and interacting with touch and spin controlled cockpits (e.g. ORF, 2012).

One way to ensure usability for cockpit interfaces and interactions is to ascertain specific user needs prospectively during the conceptual design phase of the cockpit development (Mayhew, 1999). With the usability engineering approach for cockpit development unusable cockpit innovations could be avoided. Therefore virtual test environments (e.g. CAVE) are a good alternative to traditional validation processes during the development process (Krause, 1999). Human-cockpit interactions should be evaluated virtually during different levels of the development process to ensure safety, usability and user-satisfaction. If misuses or dissatisfaction occur while interacting, rightsizing of the virtual car cockpit concept during an early stage of development process will be more easily, compared to advanced development stages with physical car cockpit prototypes (Stark, 2009). Thereby automakers will be more flexible and avoid high costs. However, prospective evaluations with physical prototypes are still underrepresented, because adaptive prototype evaluation is still time-consuming, inflexible and expensive (Wang, 2002). A logical consequence and an innovative solution would be to evaluate simple cockpit prototypes within virtual test environments. This hybrid prototyping approach allows a more realistic experience with the prototype (Stark, 2009).

Now, we are interested in the differences and explanatory variables between user performances and impressions within different test environments (modalities).

2 Research Study

2.1 Task Description and Experimental Design

In order to realize evaluations of human-cockpit interactions within virtual test environments presenting a physical interface prototype, a human-cockpit interaction set-up was simulated by using a tablet with an implemented navigation system interface, which was operated by a spin controller, within a virtual test environment (CAVE). A fully immersive five-sided 2 x 2 m CAVE system was equipped with shutter glasses to give the users stereoscopic views of the car and cockpit. Head-tracking devices mounted on the users' shutter glasses controlled the CAVE viewpoint according to the users' body and head movements. The stereoscopic application displayed a complete model of a car, including e.g. car body, seats, steering wheel, and cockpit (see figure 1).

The user tasks were to enter six predefined addresses (country/city/street) into the navigation system by interacting and manipulating the spin controller. The addresses were announced by the study manager. The navigation system was physically implemented in the virtual car cockpit, when the stereoscopic view was switched on (see figure 1). In the condition without virtual car model the navigation interface was right in front of the users chair within a white 2 x 2 m CAVE (see figure 2).

We realized a completely crossed within-design with 20 participants. The experimental design of the study had two independent variables: shutter glasses (yes/no) and modality (virtual car model vs. no virtual car model) in a within-factorial design with dependent measures of errors, task completion times and subjective data (AttrakDiff).

Twenty participants took part in the study (10 female and 10 male). The average age of the subject group was 26 years, 17 were right-handed and three were left-handed, with good stereopsis, and all with valid driver`s license. Each subject performed the task within the three test environment modalities: (A) no virtual car model; without shutter glasses, (B) no virtual car model; with shutter glasses and (C) virtual car model; with shutter glasses. We didn`t realize a complete test design, because the hypothetic modality (D) virtual car; without shutter glasses, wouldn`t make any sense.

The dependent variables were used to analyze the differences with regard to performance and impressions between the different modalities (A, B, C).

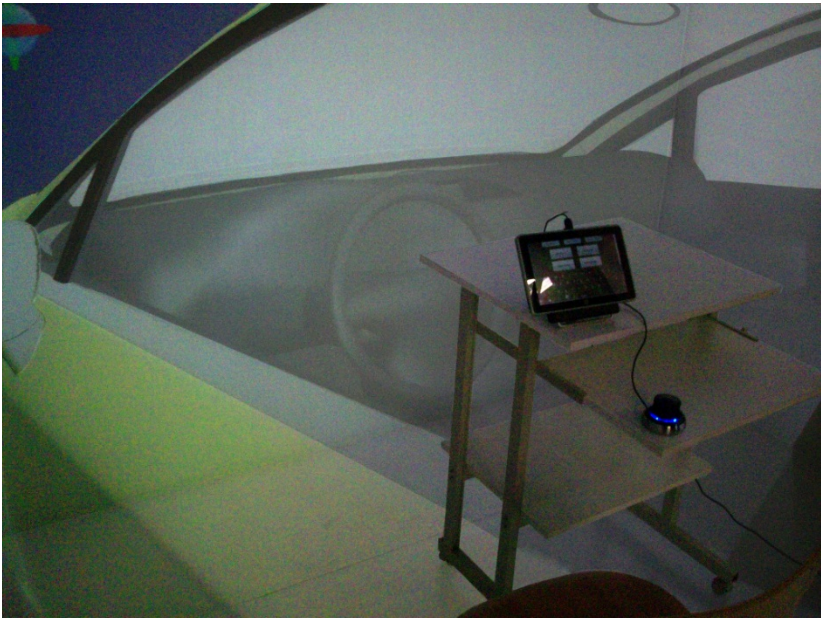

Fig. 1. Virtual car model with physical navigation interface and spin controller

Fig. 2. Participant with shutter glasses within a non-immersive test environment

2.2 Research Questions

The objective of this empirical investigation was to develop a hybrid test scenario, involving a virtual projected car model (car body and car cockpit) within a CAVE and a user interaction task with a physical interface combined with a spin controller, to answer the following research questions:

- Do shutter glasses influence users' behavior per se? Because they are slightly darkened and shutter alternating with 60Hz on each eye. The following sub-questions were asked to analyze the potential influence:
  - Do users need more time for data entries with shutter glasses?
  - Do users make more mistakes with shutter glasses?
  - Do users get different impressions with shutter glasses, compared to no shutter glasses?
- Does the virtual simulated car model influence users' behaviour? Because there will be a more realistic (immersive) test scenario within the virtual projected car, compared to a white box containing one single chair, little white table, tablet interface and spin controller. Although, users need to cognitively process more parameters of perception, majorly visual and motion parameters. The following sub-questions were asked to analyze the potential influence of the virtual car model:
  - Do users need more or less time within a virtual car model?
  - Do users make more or less mistakes within a virtual car model?
  - Do users get different impressions within a virtual car model, compared to no virtual influences?

An additional research question was prospected beside the main study intention:

- How do users rate the usability of the implemented navigation interface?

2.3 Procedure

First, the participants were required to fill out a demographic questionnaire and had to pass a stereopsis eye-test on a table outside the CAVE. After that, they were placed in the CAVE with a comfortable distance to the interface and the spin controller. Afterwards the participants had the chance to try out the navigation system by operating the spin controller.

Then the main experiment started with the first session (modality A or B). Half of the participants (n = 10) started their first trial without shutter glasses (modality A) and the other half (n = 10) with shutter glasses (modality B) within a non-immersive test environment. After completing the first session the participants were required to fill out the AttrakDiff questionnaire. Hereafter the next session within the immersive test environment (modality C) started. All participants got shutter glasses and the virtual car model was projected. They had to tackle the same task and got the AttrakDiff questionnaire once again. In the third and last trial (modality A or B) half of the participants worked within the modality they didn`t perform within the first session and reverse for the other half. The third trial was also realized within a non-immersive test environment. The main experiment was completed after two short questions (impressions about proximity to a real car).

The last ten minutes were used to fill out the usability questionnaire ISONORM about the implemented navigation system on a separate table outside the CAVE.

3 Results

3.1 Descriptive Analyses

For a descriptive and visual analysis of the results, the completion times and errors over all participants are visualized in the following table.

Table 1. Mean task completion times in seconds and average number of errors

	Modality A Without shutter glasses and no virtual car model	**Modality B** With shutter glasses and no virtual car model	**Modality C** With shutter glasses and virtual car model
Time	505 s*	499 s	451 s*
Errors	2,7	2,8	2,3

*Significant differences

The first two columns contain the times and the number of errors for the interaction task with and without shutter glasses (both without virtual car model). There is a marginal difference of 6 seconds, whereby the time is surprisingly shorter with shutter glasses. In the last column the time for the modality with a projected car model show lowest time and error rates of all three modalities.

3.2 Task Times

Task times were analyzed using a repeated measure ANOVA, with task environment (modality) as independent variable. The little difference of 6 seconds between the modalities A and B (non-immersive test environments) is not significant. It indicates that the shutter glasses do not have an influence on user performances.

With regard to the second research question, whether the virtual cockpit (immersive test environment) has an influence on user performance the repeated measure ANOVA shows a significant difference between modality B and C ($F(1,19) = 4.81$, $p < 0.005$, $n^2part = 0.202$). Hence, task completion times were faster within immersive environments, compared to non-immersive environments (see table 1), which indicates that immersive environments increase user performances. Indeed, we found undoubtful training and repetition effects. All participants were slowest in the first session and fastest in the last one.

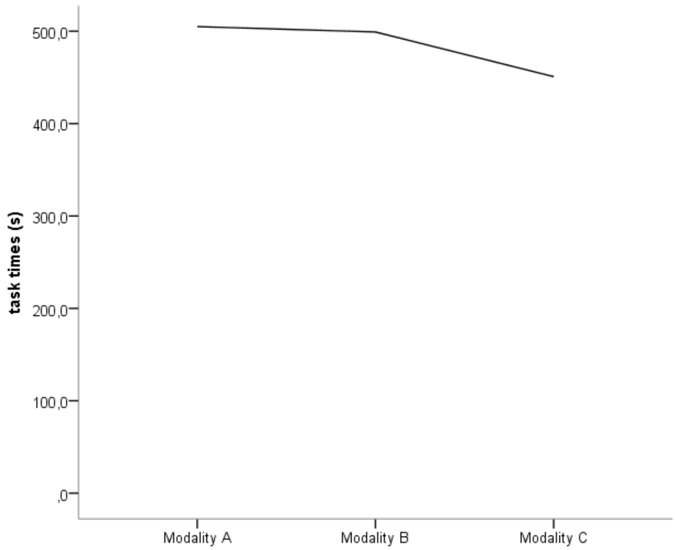

Fig. 3. Task completion times within different modalities

3.3 Task Errors

Task errors were also analyzed using a repeated measure ANOVA, with task environment (modality) as the independent variable. There were no significant differences found. But absolute mean task errors were lowest in the immersive modality C, compared to the two non-immersive conditions A and B (see table 1). This indicates a performance benefit within virtual test environment and no influence of shutter glasses, too.

3.4 Subjective Data

The AttrakDiff questionnaire (Hassenzahl, Burmester & Koller, 2003) was used to analyze subjective impressions within different test environment modalities. The AttrakDiff consists of 21 pairs of adjectives and is suitable for repeated assessments during a test session. Adjective pairs are semantic differentials rated on a seven point Likert scale. The items are grouped into three dimensions: pragmatic quality (e.g. impractical – practical), hedonic quality (e.g. tacky – stylish) and overall attractiveness (e.g. bad – good). The used AttrakDiff shows no significant differences for one of the three dimensions regarding different task environments.

3.5 Usability

To analyze the usability of the implemented navigation system the ISONORM questionnaire ISONORM 9241/10 (Prümper & Anft, 1997) was used. According to the authors, the ISONORM represents an operationalization of the seven dialog principles in ISO 9241, part 10 (ISO, 1996): suitability for the task, self-descriptiveness, controllability, conformity with user expectations, error tolerance, suitability for individualisation and suitability for learning. The ISONORM comprises five bipolar items (on a scale from -3 to +3) for each of the seven principles of ISO 9241 Part 10. In our survey we had only six dimensions (see figure 4). We eliminated the dialog principle suitability for individualization, because we decided in advance that individualization is not necessary and desired for our cockpit-interaction experiments.

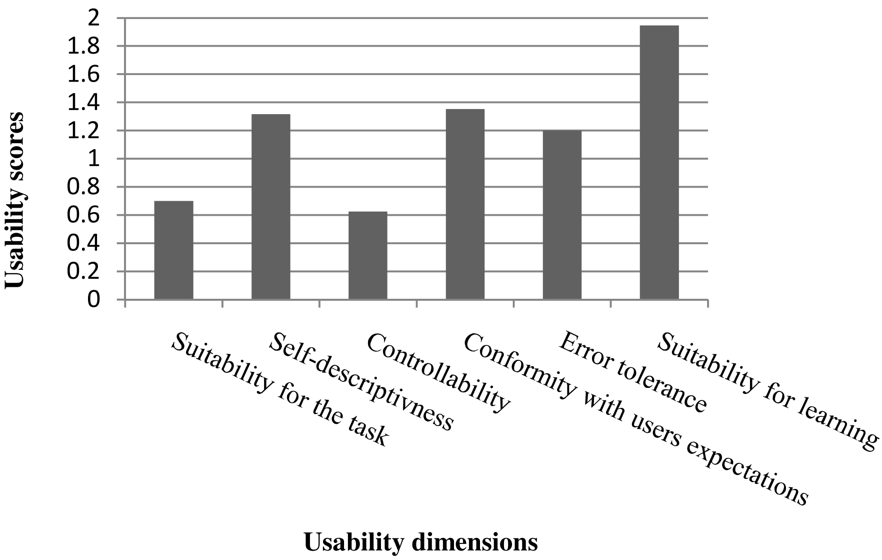

Fig. 4. Mean usability scores for all participants (ISONORM)

The intention was to detect usability issues (e.g. confusing menu structure) and/or buggy programming (e.g. non-functional menu buttons), because this navigation interface will be used for several upcoming comparative studies. The cut-off score was 0 (arithmetic mean). Figure 1 shows that all dimensions were above 0 within a positive value range. The most important dimensions for our field of application were self-descriptiveness (M = 1.3; SD = 1.2) and conformity with user expectations (M = 1.4; SD = 1.1), to avoid time-delays due to usability issues instead of modality influences.

4 Discussion and Outlook

There seems to be an explicit need for more systematic user involvement in the early stages of development processes, especially in virtual test environments. With our research study we focused on important basic parameters and evaluated their potential of influence.

Three different test environments were presented and evaluated regarding different task performances and subjective ratings: Two non-immersive modalities (with or without shutter glasses) and one immersive modality with a virtual projected car model. The aim was to analyze differences (task performance and subjective ratings) and explanatory variables (shutter glasses and virtual projected car model). To do so, different research questions were operationalized.

The first difference was analyzed regarding an important explanatory variable: the shutter glasses. The question was, if shutter glasses influence users' behavior per se, because of a slightly darkened sight and alternating shutter effects with 60Hz on each eye. There were no significant effects found for any differences between with or without shutter glasses regarding to task performance (times and errors) or subjective ratings (AttrakDiff). It is quite reliable that shutter glasses do not influence users' behavior or impressions per se. Participants do not need more time for data entries or make more mistakes with shutter glasses. Participants also do not rate different impressions with shutter glasses, compared to no shutter glasses. In summary, there are no hints to assume that limiting influences through shutter glasses will occur.

The second difference was analyzed with regard to the next important explanatory variable: the virtual projected car model. The question was, if the virtual simulated car model influences user behavior (time and errors) and subjective impressions (AttrakDiff). There are some hints that virtual simulated car models influence users' behavior positively.

Participants need significant less time within a virtual car model, compared to non-immersive conditions. Faster task completion times within immersive environments indicate that immersive environments increase user performances. Furthermore, there seems to be a tendency that participants make fewer mistakes within a virtual car model, but that effect didn`t become significant yet. Surprisingly, participants didn`t get different impressions within a virtual car model, compared to non-immersive test environments. We assume that effects will occur, when driving within hybrid test environments will be added.

After the main experiment users were asked about the usability of the navigation interface. Participants rated all of the usability dimensions from the ISONORM questionnaire within a positive range, which indicates an adequate usability value. Special focus was lying on the dimensions self-descriptiveness and conformity with user expectations, to avoid time-delays due to usability issues instead of modality influences. Both were within a good usability range. In summary, the navigation interface will be useful for several upcoming comparative studies in that field, too.

The findings presented in this paper leave much room for further research. For example, it would be interesting to see if the presented findings are applicable to different cockpit interactions (e.g. touch and gestures) or if the subjective perception differs more within a virtual driving environment compared to a conventional 2D driving environment.

References

1. Ablaßmeyer, M.: Multimodales, kontextadaptives Informationsmanagement im Automobil (Dissertation). Technische Universität München (2009)
2. Broy, V.: Benutzerzentrierte, graphische Interaktionsmetaphern für Fahrerinformationssysteme (Dissertation). Technische Universität München (2007)
3. Hassenzahl, M., Burmester, M., Koller, F.: AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualitäten. In: Szwillus, G., Ziegler, J. (eds.) Mensch & Computer 2003: Interaktion in Bewegung, pp. 187–196. B.G. Teubner, Stuttgart (2003)
4. Jahn, G., Oehme, A., Krems, J.F., Gelau, C.: Peripheral detection as a workload measure in driving: Effects of traffic complexity and route guidance system use in a driving study. Transportation Research Part F-Traffic Psychology and Behaviour 8(3), 255–275 (2005)
5. Mayhew, D.J.: The Usability Engineering Lifecycle. Morgan Kaufmann, San Francisco (1999)
6. Krause, F.-L., Franke, H.-J., Gausemeier, J.: Innovationspotenziale der Produktentwicklung. Carl Hanser Verlag, München (2007)
7. ORF Online und Teletext GmbH & Co KG, `http://salzburg.orf.at/m/news/stories/2517908/`
8. Prümper, J., Anft, M.: ISONORM 9241/10, Beurteilungsbogen auf Grundlage der Internationalen Ergonomie-Norm ISO 9241/10. Büro für Arbeits- und Organisationspsychologie, Berlin (1997)
9. Stark, R., Beckmann-Dobrev, B., Schulze, E.E., Adenauer, J., Israel, J.H.: Smart Hybrid Prototyping zur multimodalen Erlebbarkeit virtueller Prototypen innerhalb der Produktentstehung. In: Proc., vol. 8, pp. 437–443 (2009)
10. Wang, G.G.: Definition and review of virtual prototyping. Journal of Computing and Information Science in Engineering (Transactions of the ASME) 2(3), 232–236 (2002)