How to Measure UX and Usability in Today's Connected Vehicles

Rico Ludwig(✉), Arne Bachmann, Stefanie Buchholz, Kathrin Ganser, Daniel Glänzer, and Audrey Matarage

P3 Automotive GmbH, Heilbronner Straße 86, 70191 Stuttgart, Germany
{Rico.Ludwig, Arne.Bachmann, Stefanie.Buchholz, Kathrin.Ganser, Daniel.Glaenzer, Audrey.Matarage}@P3-group.com

Abstract. Besides hard factors like horsepower or consumption, today's connected cars need soft factors being available as criteria for purchasing decisions. User experience (UX), as one of these soft factors, is essentially influenced by services available inside the vehicle. Almost every new vehicle offers a variety of mostly Internet-enabled services to improve the driving experience and to make customers happier. One challenge is to make UX measurable regarding experiences and knowledge inside the vehicle. It has been researched what influencing factors significantly characterize the evaluation of in-car UX. A combination of assessment of relevance, usability testing of pre-defined use cases as well as gathering of quantitative parameters helped create a characteristic variable as a measure of UX inside vehicles. The final evaluation tool is to allow the formation of a holistic opinion about experiences and knowledge inside vehicles and to enable a comparability of UX of complete vehicles as well as of developmental states. Ultimately this not only supports customers in their purchasing decision process, but also manufacturers developing and optimizing new and existing solutions.

Keywords: User experience evaluation · Connected car · Scenario-based testing · Evaluation method

1 Introduction to User Experience and Connected Car

User Experience (UX) is an interdisciplinary field highly influenced by design, technology and psychology. In the past 20–30 years, the topic has become increasingly important. However, there is still no uniform and, above all, generally accepted definition. In the norm ISO 9142-210, for example, UX is defined as comprising "a person's perceptions and reactions that result from the actual and/or expected use of a product, system or service". This means that in order to assess the user experience, for example of a product or service, perceptions and reactions of users need to be evaluated. It is therefore not surprising that there is no general test procedure that ultimately provides a concrete value for the user experience.

Digitization in the automotive world has led to an increasing number of digital and connected services in vehicles. There are weather services, news, calendar and email

T. Ahram et al. (Eds.): IHIET 2020, AISC 1253, pp. 17–21, 2021.
https://doi.org/10.1007/978-3-030-55307-4_3

services, online radio services, emergency and breakdown services, various navigation services, music streaming services, services for chatting, gaming, telephoning, watching TV; Voice recognition, fatigue recognition, you can control your vehicle remotely or send the key digitally to friends, and of course there is a browser for surfing the Internet directly in the vehicle - in the connected vehicle, the connected car.

Whereas in the past, when car buyers were primarily interested in horsepower, engine capacity or speed, nowadays the performance of these mentioned services plays an increasingly significant role in the purchase decision. That's why today product managers and developers want to know where they stand with their digital innovations compared to their competition. These developments result in the necessity of benchmarking to make the user experience of these digital services measurable with each other.

2 How Does P3 Test the Digital Services in the Connected Car?

At P3 automotive we have been carrying out user experience benchmarks for more than six years. In 2014 we ordered several vehicles for the first time and tested them extensively. We finally published the results of our first excursion into the world of user experience testing in *connect* magazine.

This project marked the starting point in this topic. One year later we enhanced our know-how when we carried out a user experience benchmark for one of our customers.

For this UX benchmark, we carried out extensive research and identified service categories most relevant to users, such as navigation or infotainment. We then developed persona profiles as well as special use cases designed for each category and tested them live with previously acquired test persons. Since connected services are not only limited to the car HMI, our test persons had to explore all relevant touchpoints, which are defined as points of contact between a user and a system (e.g. brand). Our three touchpoints were primarily the vehicle itself, the manufacturer's smartphone app and its web portal, which was accessed on a laptop.

First the test persons should put themselves in a specific scenario where they had to complete different use cases. An exemplary use case for the infotainment category would be: display the weather at your destination. The test persons were asked to run this use case, recording their thoughts and feelings as well as their positive and negative findings. After one tester had completed all use cases for a service category and all impressions had been recorded, the overall experience was rated good/bad using a simple 7-point Likert scale. In the end, this made it possible to calculate the average value for all test subjects. This procedure was repeated for all categories and touchpoints. Finally, we let the test persons fill out the meCUE questionnaire. which is based on the CUE model by Manfred Thüring and Sascha Mahlke. It highlights the differences in the perception of task-related and non-task-related product qualities, and additionally captures user emotions [1]. In our framework it was used to enable test persons to assess the vehicle's overall experience.

Through analyzing the test results, we identified guideline values for the quality of UX for each service category. We also received detailed, qualitative feedback on the

tested services. In addition, the results of the meCUE questionnaire were another indicator for the overall comparison of the vehicles (Fig. 1).

Fig. 1. UX benchmark - P3 Connected Car Experience Days 2017

In the following years we have conducted a number of other UX benchmarks. A highlight were the P3 Connected Car Experience Days (CCED) in Berlin, when we tested the latest vehicles from ten different manufacturers and presented the results to numerous participants from the automotive industry and, above all, demonstrated them live in the vehicles. To date, we have gained a great deal of experience and identified the following relevant questions, which must be answered in advance of user experience testing for vehicles:

1. What is the goal of a UX benchmark?
2. Who are the testers and how many?
3. What test methodology is used?
4. How many vehicles should be tested?

2.1 What Is the Goal of a UX Benchmark?

The question that basically arises at the beginning of every project. Which goals should be achieved by conducting the benchmark? What insights should be gained? What types of results should be generated? It is particularly important to define which services are to be tested to what extent and at which touchpoints.

This question usually depends on the client. In marketing and sales, we now know that different insights are important than with a UX benchmark for a development department. While the UX of linking processes between driver and vehicle is

sometimes the focus, other projects require a detailed look at infotainment services such as Bluetooth audio streaming.

This needs to be clarified in advance, because the use cases will later build on these decisions as the basis of the entire benchmark.

2.2 Who Are the Testers and How Many?

A first distinction can be made between two categories: Either the UX benchmark is carried out by experts or test subjects are acquired for the tests.

If the decision is made by experts, two to four UX experts are usually enough. UX experts are fundamentally characterized by sound knowledge in the areas of usability and user experience. Extensive knowledge of the automotive sector is also mandatory for vehicle UX tests.

The situation is different for test subjects, in principle no prior knowledge is necessary and sometimes undesirable. However, a much higher number of test subjects is required. Again, the goals of the benchmark are important: If, for example, only usability weaknesses are to be discovered, good results can already be achieved with six to eight test subjects. A reliable UX test, which is supposed to provide a reliable general impression of a product or service, requires at least 20 test persons.

In order to get a more concrete picture, it still makes sense to have only certain customer groups tested and to select the test subjects beforehand.

2.3 Which Test Methodology Is Used?

The test methodology depends primarily on whether it is a test with expert or with test subjects. For the first-mentioned, heuristics or checklists are mostly used. For test subjects, however, there are numerous different methods, from simple questionnaires to Thinking Aloud to eye tracking.

Questionnaires are probably the most commonly used means of determining the quality of user experience. They are available for all aspects, including emotion detection, subjective usability and user experience, or visual aesthetics. Well-known examples are the Affect Grid, the SAM, the AttrakDiff2, the UEQ or the VisAWI. The advantages are simple usage and evaluation. However, the results are always subjective and cannot be recorded during but only after a task.

With the Thinking Aloud method, a test subject verbalizes his or her own thoughts during a test. This helps to better understand their thinking processes when working on tasks. According to Jakob Nielsen, Thinking Aloud might be the most valuable method of usability engineering [2]. It is flexible and can be used with little effort. However, Thinking Aloud can influence cognitive processes and change the behavior of the test person.

Eye tracking is particularly suitable for capturing the visual attention a test subject pays on specific areas, e.g. when selecting a radio station. The results can then be used to optimize the presentation of information [3]. Even the smallest details can be captured using eye tracking without disturbing the behavior of the test subject. However, this is associated with high acquisition costs.

In summary, it can be said that each method for evaluating the user experience in the vehicle must be individually designed with regards to the goals to be achieved.

2.4 How Many Vehicles Should Be Tested?

The question of the number of vehicles to be tested is important because other factors depend on it. First of all, the number of vehicles is decisive for the duration of the test. The more vehicles to be tested, the more time generally needs to be scheduled for the benchmark. Although this problem could be solved well by parallelization and certain study designs, the necessary number of test subjects is rarely available or only with increased resource efforts.

In addition, the number of vehicles is important for the test organization, because a sufficiently large and ideally covered location should be determined in advance.

Often the opinion between our customers and us differ on this question. Understandably, the largest possible pool of benchmark vehicles should be generated for your test vehicle. For the reasons mentioned above, however, this is often associated with considerable additional outlay or additional costs. More test vehicles may also result in a higher workload for the test persons, thus in a less reliable evaluation.

From our experience, a test set of five benchmark vehicles has proven to be optimal.

3 Conclusion

With all the experience we have gained during the last six years, we figured out that there is not the one and only UX evaluation tool or methodology that should be used as the state-of-the-art measurement tool.

Together with our clients and partners we have realized, that a set of methodologies – each one depending on particular requirements – is ideal to provide most stakeholders with their desired insights. Our set of use cases and questionnaires allows us to generate results, which are comparable between different vehicles and services and picture a clear overview and ranking.

Using additional methods like Thinking Aloud or observation ensures product owners, designer and developers to get very detailed information on how people are using their systems and software. Eventually these insights are indispensable to evolve and improve the UX of a system or service.

References

1. Smith, T.F., Waterman, M.S.: Identification of common molecular subsequences. J. Mol. Biol. **147**, 195–197 (1981)
2. Thüring, M., Mahlke, S.: Usability, aesthetics and emotions in human–technology. Int. J. Psychol. **42**, 253–264 (2007)
3. Nielsen, J.: Usability Engineering. Academic Press, London (1993)
4. Schiessl, M., Duda, S., Thölke, A., Fischer, R.: Eye tracking and its application in usability and media research. MMI-Interakt. J. **6**, 41–50 (2003)