Experiencing the future: evaluating user acceptance of automated vehicles

Yvonne Barnard, University of Leeds, UK, y.barnard@leeds.ac.uk, **Daniel de Klein,** City of Helmond, The Netherlands, **Gillian Harrison,** University of Leeds, UK

ABSTRACT

In this paper we will discuss user acceptance testing of automated driving enabled by Internet-of-Things technology, as performed in the Netherlands in the European AUTOPILOT project. The central question is: how can we test user acceptance and gather user ideas and concerns of mobility services for the future, while the technology is far from being operational and available? In other words, how can we let people experience the future in a real environment?

Keywords: user acceptance, automated vehicles, evaluation, user tests, Field Operational Test (FOT)

INTRODUCTION

1.1 Context

In the European AUTOPILOT, a three year EU-project with over 40 partners, new services were developed to enable, enhance, and accelerate autonomous driving with the aid of Internet-of-Things (IoT) technology (AUTOPILOT). In five countries tests were conducted with users, trying out new services such as obstacle detection and automated valet parking. In Brainport (cities of Helmond and Eindhoven), the Netherlands, we tested three services, with some 100 participants recruited from the general public. The services tested were: “Urban Driving”, where automated cars can be hailed that are able to detect pedestrians, and can take into account crowds of people in order to re-route, “Highway Pilot”, where automated vehicles receive information about hazards on the road, such as obstacles and potholes, and use this information to adapt driving behaviour; and “Platooning” where cars going to the same destination are brought together in order to drive together automated in a platoon on the highway. These use cases were tried out in 2019 on a test-track, a university campus and a stretch of highway.

1.2 Problems performing user tests

In an ideal situation we would let people “drive” for hours, days or weeks in an automated vehicle (AV), enhanced with IoT systems, log data during their driving and ask them about their experiences. This is the way in which Field Operational Tests (FOT) are conducted and for which a well-established methodology exists (FESTA, 2018). However, this was not at all possible. We still wanted to address user acceptance based on people's experiences. In the last few years many on-line surveys have been conducted in many countries (e.g. Becker & Axhausen, 2017). Also in AUTOPILOT we did a large survey in six countries (Aittoniemi et al., 2018a). In such questionnaires hypothetical situations are presented and people are asked whether they would find a service or an AV useful, would trust it, would pay for it etc. However, this is about systems and vehicles people have never seen, let alone experienced, or that do not even exist in real life. Now that we had vehicles available that could drive around, the AUTOPILOT

project offered us a unique opportunity to organise a real user experience.

However, we encountered many problems: *technological*, systems not always fully and correctly working; *safety related*: people not allowed to “drive”, but needing a safety driver; *environment related*: driving on a very limited stretch of road or even test track, not in real traffic, and only for a very short time; *permissions*: problems to obtain permission for the vehicles and the involvement of users; *test methodology*: wanting to ask participants too many questions but not wanting to exhaust and annoy them. In short, people only got a very limited exposure to automated driving. In the paper we will systematically address these problems.

1.3 Lessons learned and solutions

All this might lead to the conclusion that it was not a very successful evaluation. However, we found ways to overcome problems, and to create a meaningful experience for participants, and to obtain interesting ideas and opinions from them. Examples are working with story boards telling stories with pictures to set the scene, involving participants not as subjects, but asking them to envisage themselves as co-workers, focussing the questionnaires on their concerns, ideas and on future opportunities, comparing their expectations before and after the experience, and working around technology failures. Overall we got interesting results, and rather positive ones. In the paper the focus will be on the methodology of setting up and conducting such user tests. For the full results we refer to (Aittoniemi et al., 2020a and Kolarova et al., 2020). We would like to share our hands-on experiences in order to provide other researchers with ideas on how to perform these very difficult and complex tests.

USER TESTS AT BRAINPORT

2.1 Systems and scenarios

At Brainport user tests with the following systems and scenarios were performed in the spring of 2019:

Urban driving (Detection of pedestrians and cyclists, automated taxi service)

- ☐ An AV is called via the app
- ☐ The vehicle arrives at the call point
- ☐ The vehicle drives automated to destination
- ☐ Vehicle detects pedestrians (not visible, e.g. standing around a corner) and adjusts driving behaviour (stops or slows down), by picking up smartphone signals
- ☐ The vehicle detects crowds by picking up smartphone signals and adjusts its route

The Highway pilot (Hazard detection)

- ☐ Detection of road incidents and obstacles to ensure safe automated driving on highways
- ☐ Cars with sensors and roadside cameras detect obstacles, potholes, bumps, and other hazards
- ☐ Information is sent to traffic management, which determines when traffic should be informed

- ☐ Semi-automated vehicle receives a message about a hazard and adapts its driving (i.e. braking, lane-change)

Platooning (Automated (short-distance) following of vehicles for more efficient traffic and comfort) (Schmeitz et al., 2019)

- ☐ Via app, two vehicles make contact to drive in a platoon
- ☐ Both get information on the meeting point and speed advice
- ☐ When the vehicles meet, they drive in a platoon on the highway, the lead car driving manually, the following vehicle driving in automated mode
- ☐ At the destination, or in case of interfering traffic, the platoon is broken

Originally a fourth test was planned, with **automated valet parking**, but this had to be cancelled due to organisational and regulatory problems.

2.2 Study design

In Table X the participant numbers for the different scenarios are given:

Table 1 - Public Testing carried out at Brainport

	Urban Driving	**Highway Pilot**	**Platooning**
Number of participants	43	37	20
Location	Closed-off area at Eindhoven University Campus	Closed off track behind parking space Brainport Campus	Start at Brainport Campus, followed by drive on public highway

Participants from the general public were recruited via the social media of the City of Helmond, and were invited in groups of 6-8. Firstly participants were given a briefing of the AUTOPILOT project and an introduction to the technology that they were going to experience. Then the participants read and signed consent forms. They were asked to fill in a “Pre-Test” Survey which assessed their expectations about the technology and the experience. Participants were taken to the AV equipped with the IoT technology and given a short further briefing. Participants would then experience the technology in groups of 2. A professional test driver (safety driver) was behind the wheel (a regulatory safety requirement), and participants would generally be seated in the back seat. When the drive was completed the participants returned to the briefing room and completed two more surveys: a “post-test” survey to capture their reactions to the experience and future use of the technology and a “background” survey to gather socio-demographic information. Participants were then thanked for their time with a small gift, and could participate in a raffle for restaurant tickets. The whole experience lasted around 2 hours.

PROBLEMS, (SOME) SOLUTIONS AND POSITIVES

The issues we encountered in the user tests can be classified according to the phases of the test: Preparation, Performing of the tests, and the User experience.

3.1 Preparation

3.1.1 Piloting

For every FOT on a road it is clear that piloting is of the utmost importance. However, piloting means finding problems that need to be solved, the solution may generate new problems and at some point the project runs out of time, money and resources, and the real testing has to start. At Brainport we had two pilot rounds for user testing, firstly the researchers acted as participants, in the second round people working on the Brainport campus were the pilot participants. Next to a variety of problems related to the functioning of the automated systems we also found issues such as:

- Technicians and safety drivers are eager to explain the technology behind the system, discussing and explaining the systems with the participants and shifting the focus from user experience to interest in the underlying technology.
- Adaptations to the test scenarios were needed, making sure that the participants understood what was happening and where they should pay attention. For example, the Highway Pilot drive encountered several obstacles in such a quick succession that participants could not follow what the car was doing or supposed to do.
- Many small technical problems came to light. With user studies the technical side has to work nearly perfectly, there is not much value in measuring user acceptance of systems that do not work correctly.
- In the pilot it became clear that the questionnaires were far too long and questions had to be dropped. Also, as questionnaires were designed for multiple tests, countries and languages they required many adaptations.
- The questionnaire on the survey tool (Limesurvey) appeared differently on the iPad used by the participants than on the laptop with which was developed, causing some questions, and the Likert scales to be no longer comprehensible.

3.1.2 Legal and ethical issues

It was not easy to get permission for the tests from the regulators and from the ethical commission of the hosting organisation. As some test were on the public road (even the parking lot) conditions for authorisations are very strict and working with participants from the general public requires much attention to detail concerning privacy and safety; gaining approval is time consuming. For the (cancelled) valet parking test, cross-border regulatory differences caused problems with flying a drone.

3.1.3 Recruitment of participants

Recruiting participants was surprisingly easy; via the social media of the city of Helmond and via mouth to mouth publicity it was not difficult to get a sufficient number of participants. Because tests were performed on weekdays as well as weekends we got a good variety in demographics; during weekdays quite a large number of older, retired people participated. Women were under-represented.

3.1.4 Communication

A large number of parties were involved in the test, from industry, research, city, campus management, to regulator. Project partners (and equipment) came from several countries. Partners also had different backgrounds, in soft-and hardware development, user research, organisation and logistics etc. Communication between all these different partners is of the utmost importance for the understanding of the roles and expectations of each organisation and each person involved. We did encounter communication issues, but we managed to solve them in good collaboration, and the atmosphere in the evaluation team was good.

3.2 Performing of the tests

3.2.1 Technical issues

Although systems had been tested shortly before user testing began, technical issues still occurred. Sometimes these could be hid from the participants, for example, the experimenter having to manually press a button to activate a system.

3.2.2 Briefing of the participants

As participants may have been under the impression that they could “drive” themselves in a robot car, we had to manage expectations. The recruitment information was rather vague on this. In the briefing it was stressed that the systems were still under development and not a finished product. The participants were asked to see themselves as part of the project team, helping to improve the systems. We tried to create a comfortable, relaxed and open atmosphere. As most of the test situations were rather artificial, with short rides, safety drivers, low speeds, and not on public roads, we aimed to build a story around the test scenarios. In the briefing presentation we used story boards with cartoons. Even if the test scenarios could not play out the full story, it provided the participants with a wider context in which they could imagine themselves.

3.2.3 Debriefing of the participants and questionnaires

After the test users filled out questionnaires on their user experience, needs and concerns, expectations of future use, and their backgrounds. Questionnaires were filled out on iPads, which worked very well for everyone, and made data analysis easy. However, we limited the number of open questions so that people did not have to type out many of the answers. All data were kept anonymous, participants were given numbers. The contact details were only available to the city representative who was responsible for the recruitment and for the draw and could not be linked to a participant number.

3.2.4 Unexpected events

Unexpected events happen, for example, unauthorised people turning up in areas that were closed off. An indispensable technician-safety driver got ill. There was a day of very bad weather so we had to cancel some tests for safety reasons. No authorisation was given to use a parking area so a test with valet parking had to be cancelled. Because the cars/systems and technicians were only available for limited and specific times there was much pressure on getting as many tests done within a short period of time, leaving hardly any room for solving issues with the equipment.

3.3 The user experience

The participants were in general positive about their experience, although expectations before the test were sometimes higher than the appreciation after the experience. It was a disappointment for many participants that they could not “drive” themselves or were not even allowed to sit in the passenger seat. Many of the remarks to an open question afterwards were related to how interesting it was to participate in developments for the future. Some examples of remarks:

> “Nice experience, too bad we were not allowed to sit behind the wheel”
>
> “Interesting, fun to be with. Still needs a lot of work.”
>
> “Impressed and especially curious about the possibilities and future applications”
>
> “I was a bit disappointed. My expectations were set too high.”
>
> “I did not feel unsafe or uncomfortable but I had a few seconds of uncertainty and doubt as the car braked in front of the pedestrian. The camera screen made it clear that they had been seen by the car but I was thinking about situations where the obstacle is less obvious (kids or animals, branches on the road) and was missing a signal from the car to make sure the obstacle has been perceived.”

We asked the same questions about user attitudes before and after the drives with the AVs. Figure 1 provides the results for the platooning as an example. Users were asked what they expected the experience would be, and what their experience of the test was on a five point Likert scale (2 positive, -2 0 neutral and -2 negative) (Van der Laan et al., 1997). This figure shows that experience can make a difference. For example, it was more boring than expected, but not more dangerous.

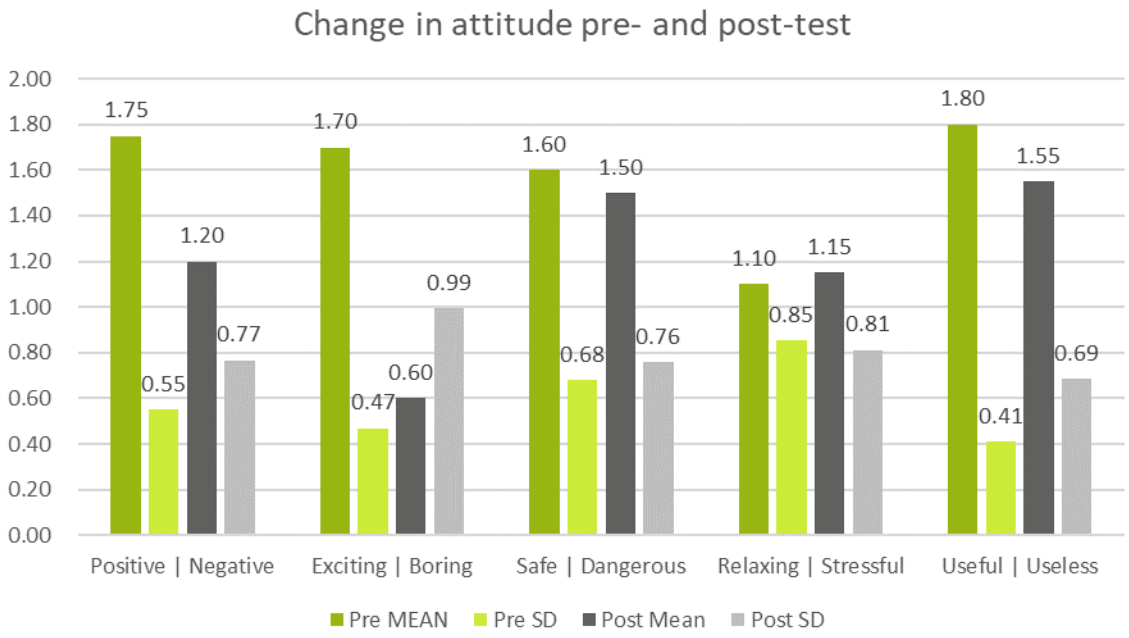

Figure 1 – Attitudes before and after test drive, for Platooning scenario

LESSONS LEARNED

Performing user tests with complex systems such as automated vehicles in a natural environment are without doubt complex, and problems can be expected on all levels. However, in the FOTs and pilots with automated systems there are three major categories of questions to be answered

1. What is it that we want to find out?
2. How do we test systems that are still under development?
3. How do we give users an experience that is relevant for getting answers on future use and user acceptance?

4.1 What is it that we want to find out?

In the FESTA methodology for designing and performing FOTs, research questions play a central role in the study design (FESTA). Research questions determine what the FOT wants to find out, and how the study design will be able to answer these questions. Although research can be related to how the system will behave in a variety of real-life conditions, research questions also address the reaction of the users. In the large European projects defining the research questions has always been a difficult process. There is so much we want to find out, stakeholders are interested in different questions, and budget and time are always limited. Defining research questions is usually not such a problem, but selecting which questions to address is. In the AUTOPILOT project this was also the case. For the user acceptance work it meant starting with a large set of questions, as already mentioned, and these were to cover many use cases in many countries, adding to the complexity (Netten et al., 2018). Even after

bringing down the number of questions we found in the pilot that the questionnaire was still too long for users to fill in comfortably. Having too many questions could lead to participants becoming annoyed and no longer paying much attention to the questions and thinking about the answers, so the data become less useful.

There is no easy way out as restricting the number of research questions in order to restrict the number of questions to ask to users, means disappointing some stakeholders or researchers with specific interests. For example, should we try to find out how the experienced systems may change users' future mobility or how much they would be willing to pay for such a system? Also the system developers have very specific questions for their system. In addressing this problem we would recommend the following guideline: which questions can only be answered by experiencing the technology, such as a drive in an AV (or with some automated system) and which questions could be answered without. Or to be more precise, what is unique about this technology than can only be understood by experiencing it?

4.2 How do we test with systems that are still under development?

The FESTA methodology was developed with systems in mind that were relatively mature, near market. However, in the AUTOPILOT project development of the prototypes was still ongoing. This meant that there were many technical difficulties, and that the systems with which we performed the pilots were still being improved. Having limited time for the vehicles being available to perform tests, specialised personal needed to operate them, coming from different countries, and time needed to obtain permissions made user testing a logistic challenge. In addition, infrastructure needed to be installed, such as cameras, test tracks fenced off, potholes created (for hazard detection), security guards hired, etc. etc. All this meant that only a limited number of days was available in which as many tests as possible had to be crammed. Improvisation was needed when things did not work according to plan. However, exposing participants to experiences with vehicles in a sometimes artificial environment, with vehicles at a low speed, a safety driver, and with user interfaces that were not made for usability, meant having to explain how the system should work in "reality" in the future, appealing to the imagination of the participant instead of just having the experience.

We would recommend to provide more time and budget for these studies, but that is of course not very realistic. What we could recommend is to allow for more separation between development and user testing, and having a clear cut-off date and go/no-go decision on whether to proceed with user testing or not. Also attention should be paid to the communication between partners, to ensure true collaboration between developers and evaluators, for example, by establishing clear lines of contact and responsibilities. Originally we had hoped that the testing with people working on campus could have been the first real test, but because of the many difficulties encountered we decided to treat it as a pilot test. Another way of looking at this dilemma is to decide not to worry too much about the technical and logistical difficulties, and to try out whatever is possible, just to get in information. However, these tests are very costly, so it is important to make decisions, and preferably early on, on how to use and divide the resources available in a project. We did underestimate the costs involved in performing the

evaluation.

4.3 How do we give users an experience that is relevant for getting answers on future use and user acceptance?

Given that we are looking for added value of user experience above just eliciting opinions via on-line questionnaires, we need to make sure to provide participants with a relevant experience. Given the technical and logistical problems described above, this is not an easy requirement. One of the biggest disappointments for the participants was that they were not allowed to sit behind the steering wheel themselves.

Although more and more real users tests are being performed, specifically with automated shuttles providing normal services, testing advanced systems such as the ones used in AUTOPILOT encounters issues with having artificial environments, artificial scenarios, safety drivers, low speeds, user interfaces that are not fully developed or not easy to understand, and (very) short time space for users in the vehicle. Use of videos can help to provide participants with a better understanding of the capabilities of the systems, but experiencing being in an AV is something different.

By taking several measures we aimed to improve the user experience, starting with managing expectations:

- Explaining we saw them as team members, not as potential customers, in order to encourage them to be critical
- Creating an informal, pleasant and relaxed atmosphere
- Briefing the participants using a story, in the form of a cartoon, with recognisable characters, such as the story of Bart and Wendy who arranged a platooning drive. The briefing also explained how the system would function in future services.
- Explaining beforehand that not everything would be perfect.
- Avoiding detailed technical explanations and questions to the technicians and safety drivers, aiming to keep the participants in the role of users, stressing that the organisers of the tests were not involved in the construction of the systems

By taking care that participants enjoyed the whole test event, even if they did not like the system itself, also ensured that they would not cause negative publicity about the project or the organisation responsible or involved in the user testing.

Concluding remarks

In conclusion, lessons-learned and recommendations for user testing of AVs and systems are partly overlapping with those for user acceptance) and user experience testing of technology in general (e.g.

Venkatesh et al., 2003). What is specific for the user testing of AVs is the large gap between expectations and promises of the vehicle automation, and the actual stage of development. Furthermore, vehicle automation impacts many areas of the future mobility of people, as well as the liveability of cities and towns (Aittoniemi et al., 2018b & 2020b). We are asking people to project themselves into the future, they are not only asked to think whether this is a useful and convenient product for them, but to imagine how their daily life and their environment could change. As researchers, industry and cities we may have our own ideas what a future with full road automation may look like, but we do not actually know, and there is a range of completely different future scenarios. We do need tests in which members of the general public engage in trying out new developments, providing their ideas and concerns, and help shape that future.

ACKNOWLEDGEMENTS

The project “AUTOPILOT – Automated driving progressed by Internet of Things” was funded by the European Union´s Horizon 2020 Research and Innovation programme under Grant Agreement No 731993. The authors want to thank all project partners who contributed to developing, organising and performing the user tests, specifically the City of Helmond, TNO, and Eindhoven Technical University, and the partners responsible for the technical development and trials of the systems tested. Finally we would like to thank all the participants who gave us their time, ideas, and their good company, and made the experience a very enjoyable one for the evaluation team.

REFERENCES

Aittoniemi, E., Barnard, Y., Ertl, D., Harrison, G., Kolarova, V., Malin, F., & Touliou, K. (2020a). User requirements analysis. AUTOPILOT Deliverable D4.8. FIA. Available at: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5cce7a6d9&appId=PPGMS

Aittoniemi, E., Barnard, Y., Harrison, G., Innamaa, S., Malin, F., Rämä, P. (2020b). Towards a method for getting a grip on societal impacts of automated driving. In: Proceedings of 8th Transport Research Arena TRA 2020.

Aittoniemi, E., Barnard, Y., Chen, H., Ertl, D., Harrison, G., Kolarova, V., Malin, F., & Touliou, K. (2018a). User requirements analysis. AUTOPILOT Deliverable D4.7 User requirements analysis. Available at: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5cce7d663&appId=PPGMS

Aittoniemi, E., Kolarova, V., Barnard, Y., Touliou, K., & Netten, B. (2018b). How may connected automated driving improve quality of life? Paper presented at 25th ITS World Congress, Copenhagen, Denmark.

AUTOPILOT, AUTOPILOT Project Brainport Pilot Site: https://autopilot-project.eu/brainport-2

Becker, F., Axhausen, K.W. (2017). Literature review on surveys investigating the acceptance of automated vehicles. Transportation 44, 1293–1306.

FESTA (2018). FESTA Handbook, version 7. Available at: https://www.connectedautomateddriving.eu/methodology/festa/

Kolarova, V., Ertl, D., Aittoniemi, E., Harrison, G., Touliou, K., Barnard, Y., & Naendrup-Poell, L. (2020). Assessing user expectations, requirements, and concerns toward automated driving progressed by internet of things – a user-centric development approach. Proceedings of 8th Transport Research Arena TRA 2020.

Netten, B. (Editor), Aittoniemi, E., Barnard, Y., Federley, M., Gaitanidou, L., Karagiannis, G., Kolarova, V., Malin, F., Lenz, O., Netten, B., Pont, J., Van den Boom, B., & Willenbrock, R. (2018). AUTOPILOT Deliverable D.4.1 Methodology for Evaluation. Available at: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5b83669f7&appId=PPGMS

Schmeitz, A., Schwartz, R.S., Ravesteijn, D., Verhaeg, G., Altgassen, D., & Wedemeijer, H. (2019). EU AUTOPILOT project: Platooning use case in Brainport. Paper presented at 13th ITS European Congress, Brainport, the Netherlands, 3-6 June 2019. Available at: https://autopilot-project.eu/wp-content/uploads/sites/3/2019/05/ITS19-AUTOPILOT-Platooning-TNO-FINAL.pdf

Van der Laan, J.D., Heino, A., & De Waard, D. (1997). A simple procedure for the assessment of acceptance of advanced transport telematics. Transportation Research - Part C: Emerging Technologies, 5, 1-10.

Venkatesh, V., Morris, M. G., Davis, G. B., and Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS quarterly, 425-478.