BMW GROUP

Design and implementation of a strategy that introduces customer feedback in the early pre-development phase of Natural User Interactions in BMW car products.

Gilles Ouwens

Colophon

BMW GROUP

Author
Gilles Ouwens

Master's thesis
Msc Strategic Product Design
Industrial Design Engineering faculty
Delft University of Technology

Graduation Committee
Chair - Dr. ir. E. A. van den Hende
Faculty of Industrial Design Engineering - Assistant professor of Innovation Acceptance and Consumer Research

Mentor - MSc. M.J.J. Buijs
Faculty of Industrial Design Engineering - Course coördinator of Build Your Startup

Company mentor - MSc. S.E.H. Heijboer
BMW AG - UI/UX concept designer

Illustrations - BMW AG

Abstract

Hereby, you are reading my graduation project for the Master's program in Strategic Product Design on the TUDelft. The research has been formulated from my interest in both customer-centric design and the automotive industry.

This research started with the aim to introduce short-cycle feedback loops in pre-development departments to enable customer feedback in the early development stages to enable designers and developers to validate their product concept. The departments that will be targeted are the UI/UX pe-development departments of BMW, to be exact teams and departments that work on Natural User Interactions. The BMW Group introduced the term BMW Natural User Interaction (NUI)*, at the Mobile World Congress 2019 in Barcelona. The NUI combines advanced voice command technology, gesture control and gaze recognition to enable more real-world interaction between the user and product.
Earlier gained knowledge and insights enabled me to see many benefits of implementing customer feedback as early as possible. To analyze the possibilities of customer feedback, literature research was combined with a company analysis. From an outsider's perspective, it seemed that the automotive industry is a bit careful and hesitant to involve customers in development processes. Therefore, the motivation of this industry could still benefit from novel customer feedback implementations. Both the customer side and the business side, benefit from streamlined customer feedback implementation in product development. Therefore, this research aims to validate if customers would sense an ability to influence the end products of BMW.

The company analysis started from researching how departments function and interact with each other, to their development processes and finally the data gathering analysis. From the company analysis, the assumption of potential for early customer feedback was validated.
From this point onwards a search for suitable customer feedback methods was started, after considering and analysing many combinations from contemporary methods with new user interaction technologies, a 3-step method was designed to test my proposed VUI IN-Car method.

I am proud to say that promising results have been found from the 3-step test method. Interacting with a speech interaction while giving feedback on speech functions, results in a higher amount of creative ideas from participants. And adding a more realistic context to this setup results in more creative input. These results were compared to a contemporary questionnaire with the same content. The detailed steps I took will be explained in this report.

Table of content

BMW GROUP

BMW MINI

1 - Introduction

1 - Introduction

Stakeholder information

To begin, this research was executed for both the BMW Group and the TUDelft. Since, projects can be copied by the competition, sensitive information, and in the end, detailed information in the Strategy for BMW, are being held separately from my thesis. My work for the TUDelft contains my methods, customer testing and strategic decisions that resulted in strategic advice. Most of the information from BMW has been available to my Chair and Mentor but is documented separately. My strategic report will not incorporate the tailored BMW advice that is presented to BMW separately.

Personal Motivation

The foundation of this Master's thesis started from a combination of personal experiences and interests. I developed an interest in the relationship between customer and brand: How does a customer think, and how do brand and customer influence each other? Figure 1 Ingredients for better interaction In my Strategic Product Design Master's, I learned how to analyze companies and create strategic solutions for them. From a 3D Automotive Design course, I got inspired by VR-Design technologies. After combining my experiences and interests, I developed a target; new technologies can and should be used to enable improved interaction between car brands and their customers.

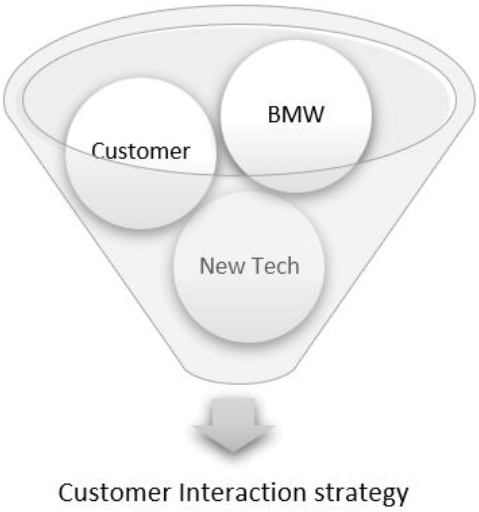


Figure 1 Ingredients for customer interaction

MainGoal

In short, the main goal of this research can be described by: Introducing short-cycle feedback loops in pre-development departments to enable customer feedback in the early development stages of Natural User Interaction products, in order to enable designers and developers to validate their product concepts or visions in a very early development stage. BMW Group introduced the term BMW Natural User Interaction (NUI), at the Mobile World Congress 2019 in Barcelona. The UI combines advanced voice command technology, gesture control and gaze recognition to enable more real-world interaction between the user and the product.

Internally, these customer feedback loops provide early iteration rounds and empower developers and designers to present their concepts or visions to their responsible superiors with proof of concept, supported by consumer feedback results. Externally, these extra touchpoints between the customer and Research & Development will contribute in a way that Hair et al. (2016) described by that; "customers who are empowered to create, report a more favourable attitude towards the brand, enjoy using the co-created product more and are more willing to spread positive word-of-mouth as well as to pay more." Therefore, this research aims to validate if customers would sense a high level of co-creation or ability to influence the end product of BMW.

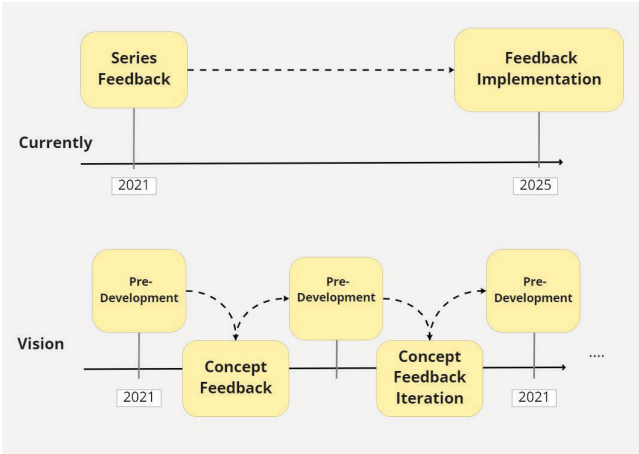


Figure 2 Short loop

Approach

To analyze the current way of customer feedback, the literature research is being combined with the current approach BMW has to gather feedback from customers. Currently, BMW departments mainly hold user studies and the end of their development phase. The departments executing these studies are open to giving insights into this research and showcasing the current way of gaining customer feedback. Findings from the literature study were combined with results from the BMW analysis and were categorized as for example; `Company structure´, `Opportunities´ and `Methods´ to make customer feedback possible or a summary of current methods. These findings lead to new `User tests´ and `Tool sets´ that are tested in the ´Concept phase´ of this research. From these results, the research questions were formulated. Below a summary of the research process is illustrated in figure 3.

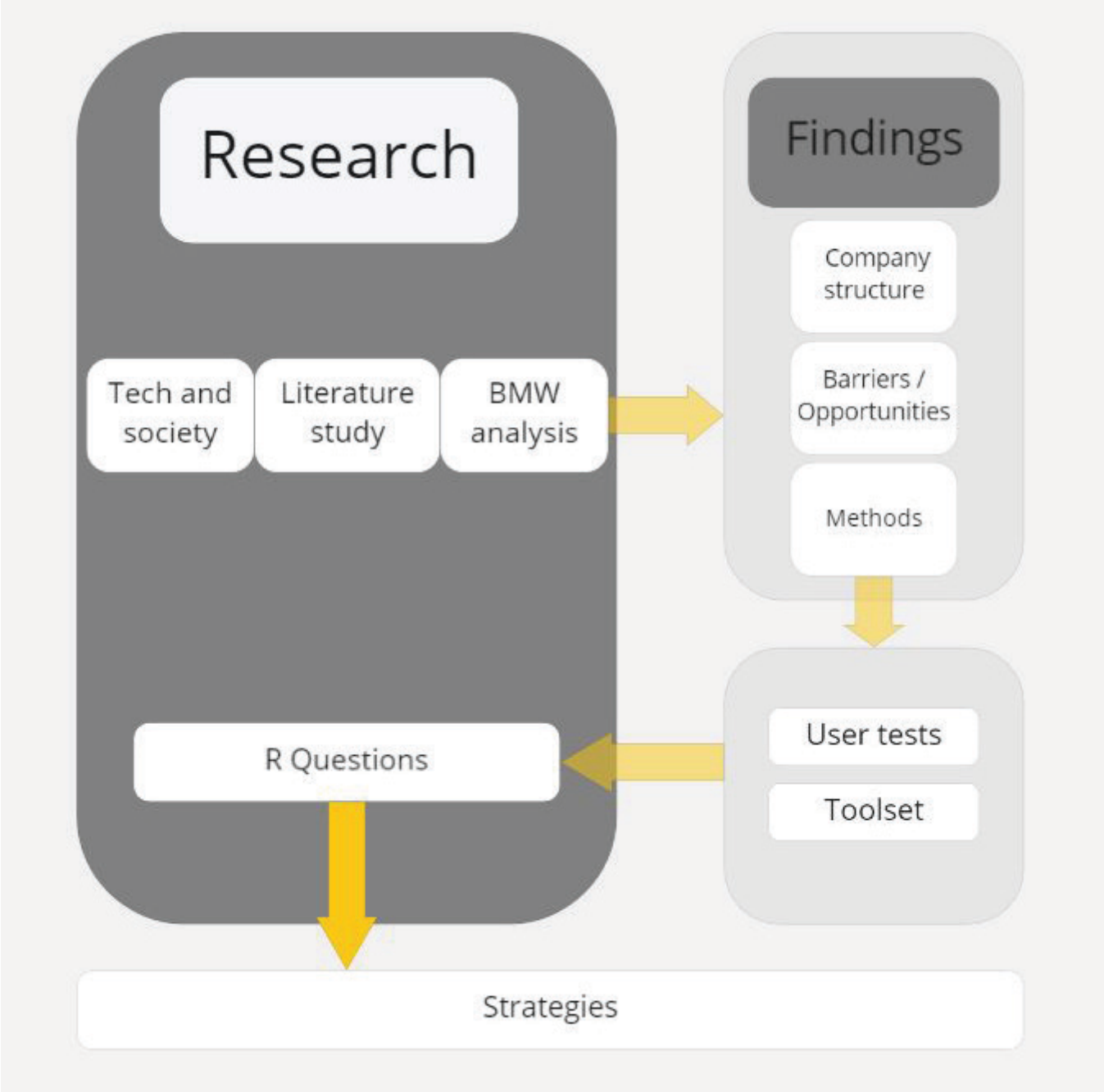


Figure 3 Approach

Build a framework

In order to build a framework, my literature study and initial BMW employee interview research consist of sub-research questions. These questions supported the research focus and consisted of questions that answer the following topics:

-	Type of departments;	What departments benefit from customer feedback?
-	Type of feedback methods;	What type of customer tests have potential?
-	Type of feedback moments;	When is customer feedback needed?
-	Type of end-customers;	Which customers are needed?
-	Type of data;	What customer feedback is desirable?

Tools

The group of technologies that are part of BMW's pre-development phases were considered in `Analysis 2´ as possible tools and therefore considered to be part of the `Concept phase´. The following technologic tools were available from BMW:

VR technology – Virtual Reality goggles and supporting expertis
AR technology – Augmented Reality prototype technologies and supporting experts
VUI – Prototyping software ´Voiceflow´ enables Voice User Interactions (VUI)
HUI – Haptic screens with prototype software to simulate Haptic User Interactions (HUI)
GUI – Gesture sensors and prototype hardware to simulate Gesture User Interactions (GUI)

2 - Analysis I

2 - Analysis I

Literature Research

My Initial literature research is divided into two categories of papers, with subjects related to:

- Co-creation in design/new product development
- Psychological design / How to test with end-users

The most important finding from this literature study is that customer co-creation is regarded as a great solution towards better design results, increases customer bonding and raises willingness from customers to pay more. By researching how to properly gain customer feedback, topics such as biases and preventing design fixation contributed to the execution of my customer tests.

Customer co-creation

One of the main reasoning behind testing with customers is not to rely on internal expert knowledge alone. Considering the principle of cognitive dissonance, it is difficult to critically assess a concept or product where you are deeply involved (Carbon, 2015). An unbiased view is for that reason a desirable factor to avoid loyalty conflicts to a concept or development.
In design-driven innovation, co-creation or customer involvement can be used as a tool to validate product concepts or product visions. Prototypes are often the tool to test a product to customers, however, in early pre-development phases, there are mostly no prototypes ready to be tested. Developers in this stage think they are not ready for consumer testing and postpone testing with customers to the development phase. Therefore, it is seen as a challenge to explain new concepts to a customer, because of the unavailability of testing with prototypes (van den Hende et al., 2012). Decisionmaking is a big topic in pre-development phases, and therefore involving the influence of the consumer can have great positive effects in this phase.

Customer Bonding

The beneficial effects of co-creation do work in two directions. Since studies have shown that customer involvement in value co-creation affects customer satisfaction in a positive way (Hunt et al., 2012). Not only the product development side benefits from co-creation but customers feel they have a say in the end result and therefore bond more to a brand or product. One can conclude that by incorporating customer co-creation the connection between customer and product can be increased.

Understanding Humans in Design

My attraction to product design is to create for humans, therefore I feel the need to understand them. To my belief, the Psychological understanding of humans makes you a better person and designer. Like Carbon states, "Without psychology we will not understand what is going on in humans, why they want to use or avoid using a product, why they admire or hate products and why they fail to use some routines or feel pleasure or discomfort using them."(Carbon, 2019).
However, just retrieving information from a customer to understand what they prefer or desire is not so straightforward. A TUDelft teacher once told me, "If you ask a customer what car they want, they will answer you by saying something like; my previous one but then newer and better". Customers normally prefer the known and trusted solutions, therefore will not prefer innovation. "This conservative tendency is natural, as familiar products don´t require any extra learning compared to innovative solutions."(Carbon, 2015). Another important remark is that "Projects might become less innovative with input from end-users (Gourville, 2006; Tauber, 1974). This originates from the fact that customers generally have a limited understanding of product innovations and are uncertain about the new benefits

(Christensen, 2007; Gourville, 2006; Tauber, 1974).
And it seems that regular new product tests do not provide customers with suitable information to let them ignore their natural resistance to innovative products (Lynn et al., 1996).

Design Fixation

Like nature, people tend to go towards the path of least resistance, this refers to the cognitive tendency to make the least effort possible when dealing with creative tasks (Ward, 1994). As a result of this tendency, we need to be careful in helping people to perform creative tasks. When people examine examples these may unconsciously stimulate them to create variations of the examples, instead of creating new products or ideas (Cheng et al., 2014). In design, this phenomenon of least resistance is called design fixation, and this is what we need to prevent from happening to extract valuable customer feedback.

Solutions

Luckily, humans still feel attracted to innovations. But in order to accept innovative concepts, we need some time to understand new and unknown products. In this belief, two possible solutions were explored to exclude the discussed possible pitfalls by asking customers to give feedback. The two possible solutions that have been explored can be seen as a form of giving examples to individuals.

Giving Examples

When giving individuals the task to generate design solutions, a study showed that giving examples to those individuals, enables these individuals to dive more deeply and narrowly into the problem. This results in higher quality and novelty of their output (Sio et al., 2015). By providing a single example research shows that the results became even more focused. However, an uncommon example will trigger people to create solutions regarded as less typical. "Together a focused search in an uncommon domain should facilitate novel conceptual combinations (Sio et al., 2015). According to this theory, I concluded that giving an uncommon single example of a concept or new product gives focus and novel ideas. In order to stimulate creative solutions from people, there are several techniques to give examples. Two of these techniques are explained below.

RET

As mentioned by giving examples we stimulate the output of individuals that need to create solutions. And to gather valuable insights from customer feedback, "a type of "systematic familiarization" is needed in order to obtain valid judgements from typical consumers." (Carbon, 2015). Therefore, the `Repeated Evaluation Technique´ (RET) was developed to familiarize customers with products in a targeted way, to replace a familiarization phase for innovations or new products (Carbon, 2005). RET shows a simulation of future possibilities and lets customers think in a targeted and intensive manner about the concept that it enabled them to thereafter make specific judgements. RET mainly contributes to long innovation cycles, therefore very suitable in the automotive industry, where product corrections take too much time and money to correct. RET was considered to be a possibility to fit my solution as "using RET, you can recognize the first tendencies for low acceptance already at the pre-development, development and pre-market launch phases."(Carbon, 2015). As stated in the example chapter, giving one uncommon example would generate focus and novel ideas. By combining this with the RET technique, I considered it to be beneficial to repeat this uncommon example to facilitate "systematic familiarization".

Narrative transportation as a solution

"Narrative transportation is a mixture of attention, imagery, and feelings that people experience when they watch a movie or read a narrative" (van den Hende, 2012). The comparison with being part of a movie or storyline instead of reading a storyline seems fitting. "Prior research on mental visualization and really new products by Hoeffler (2003) show that mental visualization instructions help customers learn about discontinuous new products that they cannot compare to existing products." (van den Hende, 2012). Therefore, narrative transportation is considered to be a solution to help the customer learn about totally new concepts or developments.

"Once customers have imagined and experienced the technology application through narrative transportation, the surrogate experience will feel real and compensate for the lack of realism in the (visual) information that is provided. In addition, the technology application can be represented with less realism (i.e., as a drawing) as long as the application is explained in a narrative form" (van den Hende, 2012). This explanation taught me that the goal of using low fidelity customer tests to generate creative data from customers is justifiable to investigate. Hence, customer tests without high-end prototypes require less financial input and are also less likely to give away sensitive information about new developments. Because an abstract representation of the new technology is being tested in a customer test application.

Concluding

Comparing the two discussed solutions; giving examples to inspire individuals by familiarizing individuals like the RET method states, or using narrative transportation that learns customers about new concepts, we can see the common aspect in helping the customer to activate their creative tasks. Making use of either of these two methods to prevent design fixation seems to work around the natural tendency of people to go for the path of least resistance. This knowledge formed my proposed method in the `Concept Phase´, by first informing and thereafter questioning the customer to give feedback.

Analysing BMW

To understand when and how short-cycle customer feedback will be beneficial, thorough research about the BMW organization was needed. Starting with analyzing the departments where development processes are being executed. In this analysis, the desirability of short-cycle customer feedback was tested. Thereafter, an analysis of contemporary processes and an analysis of the desired data will be made. From this knowledge, a systematic approach will be designed that will validate if a new method can provide short-cycle consumer-feedback loops.

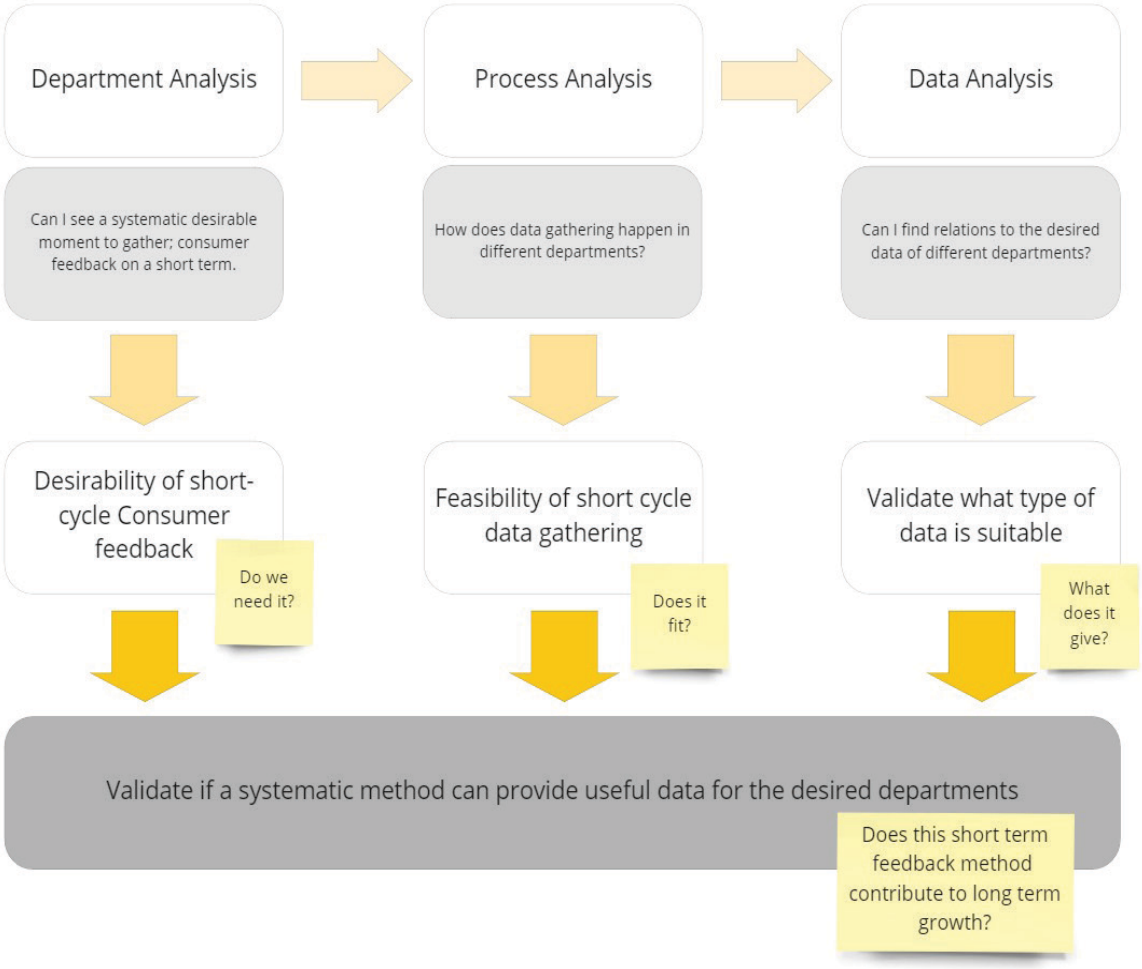


Figure 4 Company analysis

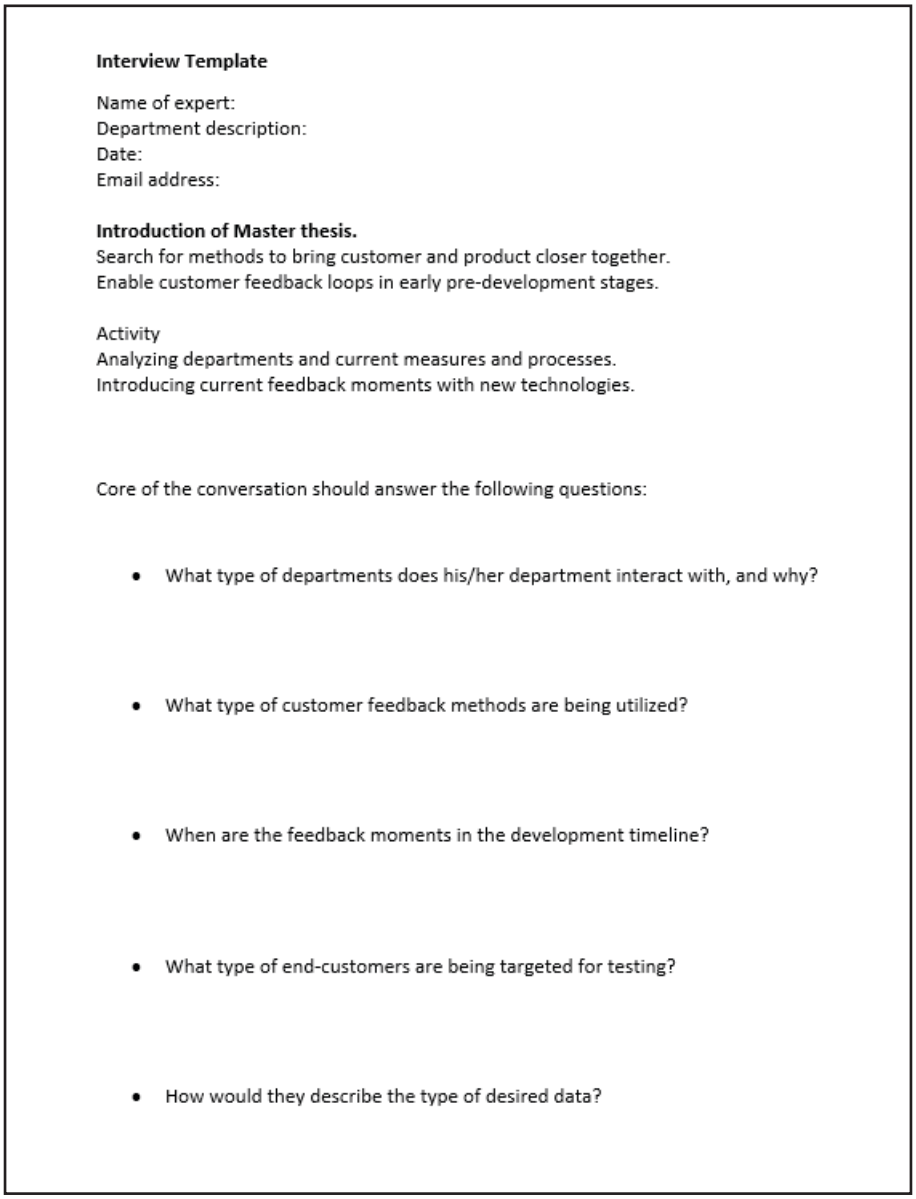
Expert Interviews

The chosen way to understand this organization is to perform interviews, this provides valuable insights and will enlarge the understanding of the company context (van Boeijen et all., 2013). To target the interviewing phase a selection has been made of BMW experts to enable the company analysis and later decision-making. In table 1, an overview of the experts is shown. Research has shown that for consumer needs 10 to 15 interviews will reveal about 80% of the needs (van Boeijen et all., 2013). However this research is analysing the needs of a big organization, therefore a higher amount of interviews was expected to be desirable and took place for that reason.

Expert Interviews	
#	Expertise
1	Haptics UI/UX
2	UI research
3	Series implementation
4	Later user testing
5	Consumer Analitics
6	UI/UX passenger entertainment
7	BMW customer databank
8	IPA prototyping (China), customer feedback
9	UI research, bridge apps protopie tanvas display
10	Working student Skylab UI Prototyping
11	Product Strategie. Customer Insights Manager My Journey
12	Steeringwheel IX
13	Autonomous dealership UC, digital aftersales PHD
14	Market Research, Early feedback tests.
15	Internal customer tests
16	PO Voice interaction
17	Produktportfolio, Roadmap
18	BMW Story Uden / ´s Hertogenbosch
19	BMW v. Poelgeest Amsterdam
20	BMW de Beier Heereveen
21	UX-designer early phase 2025+
22	User Speech testing
23	User assisting interaction concept development
24	UI User functions
25	Partner user functions
26	Product Owner Customer Entertainment Infotainment
27	Head of Groundconcept User Interaction
28	Practice Lead User Interaction
29	UI PO Practice Lead
30	Usability Studie facilitator
31	User assistent prototyping (China), customer feedback

Table 1 Expert list

In order to gain comparable insights between the interviews of employees from a wide range of departments, a semi-structured interview template has been made for all BMW experts and department employees. This template makes sure the interviews gain information about department procedures, process steps and data gathering.



Interview Template

Name of expert:
Department description:
Date:
Email address:

Introduction of Master thesis.
Search for methods to bring customer and product closer together.
Enable customer feedback loops in early pre-development stages.

Activity
Analyzing departments and current measures and processes.
Introducing current feedback moments with new technologies.

Core of the conversation should answer the following questions:

- What type of departments does his/her department interact with, and why?
- What type of customer feedback methods are being utilized?
- When are the feedback moments in the development timeline?
- What type of end-customers are being targeted for testing?
- How would they describe the type of desired data?

Figure 5 Interview template

Department Analysis

To understand the core of BMW's organization and the individual departments, interviews with internal experts and employees were being held about the company structure and how internal departments work together to achieve their goals. In order to work towards solutions, it was crucial to understanding the goals and content of departments. In this phase of getting to know how departments function that are involved in customer feedback, the first step was interviewing the market research department. There the insight was gained that currently, user tests take a long time in the preparing phase, "..the developers come to us with a question to do a study. Then the time and money need to be asked to set up this study, this takes a few months. Then, the next steps are taken to really set up the study" (Expert, 5). From the product development side, another viewpoint gave the following insight: "You are completely right that

we should validate our concepts more and earlier with real customers" (Expert, 16). This does reflect my assumption that short-loop feedback implementation could add value to this organization. From the interviews, have been concluded that there is a lack of early customer feedback and indeed a desire to validate concepts or visions in pre-development stages.

Process Analysis

In order to create new customer feedback possibilities, analysis was needed of current interaction moments between BMW and their customers. First, to clarify the current measures that are being used to receive feedback from customers. From this analysis, I concluded that applying user tests is the main method that currently is being applied by BMW to collect customer feedback. These user tests tend to use high fidelity prototypes in order to give the customer a realistic experience of what the product could be.

Next to well-prepared user tests with actual BMW customers, there are shorter tests in the series development phase with internal BMW employees. These employees are being considered `internal´ when they work for BMW directly and are not hired via an external supplier. "Our software tool enables a pre-selection of the participants by giving the possibility to exclude participants by their departments or their expertise or knowledge" (Expert, 30). By making use of this pre-selecting tool BMW simulates the use of experienced customers, without exposing real customers to classified information. As literature states, a solution to prevent design fixation on examples in customer studies could be; to include information from experienced lead users in the NPD process (Schoormans, Ortt, and De Bont, 1995; Von Hippel, 1986). This, therefore, was considered to be a grounded source of gathering feedback. In order to prevent biases as much as possible, a pre-selection of the BMW employees has been made. Only employees from departments outside the innovation- and design departments were allowed to take part in the research tests. In this way employees with related knowledge and therefore biased employees can be excluded from this research.

Next to the insight that currently most user tests make use of realistic experiences, I concluded that there are tools to test with internal employees to simulate experienced users, this would be a suitable tool to use to test my methods later on.

The contemporary main user tests that are being used to gather customer feedback in the BMW pre-development phase are called 'Late User tests' in this research and are illustrated by the red squares in figure 6. "These tests are being held near the end of the pre-development phase, where last iterations of product concepts and product features are tested in order to finalize them" (Expert, 11). "These [tests] are too late for testing requirements [they] are for, developers that know: at this point, we have a problem and we have to solve this and this…"(Expert, 24). These late user tests are used to solve known problems. Afterwards, the concepts are taken over by the departments that are part of the successive phase called `Serie Development' in the `Timeline user tests´ figure. "It is true that there is limited time left to make iterations based on the results from these tests" (Expert, 7). The illustrated `Early user tests´, indicated in yellow, are short iteration customer tests where development projects can compete to test their concepts. "These tests are not being used that often because few see the value gain compared against the effort of organizing a user study and analyzing the result afterwards. Currently, developers rather choose to rely on expert knowledge instead" (Expert, 16). Concluding from expert interviews "there are two moments where short customer feedback loops could add beneficial value, very early requirement testing and internal user test" (Expert, 15). The suggested moments are very early in the pre-development phase and in early series-development to finetune and confirm the implementation of features.

These moments are illustrated by the small green squares called `Concept aim´.

- Concept aim
- Early user tests
- Late user tests

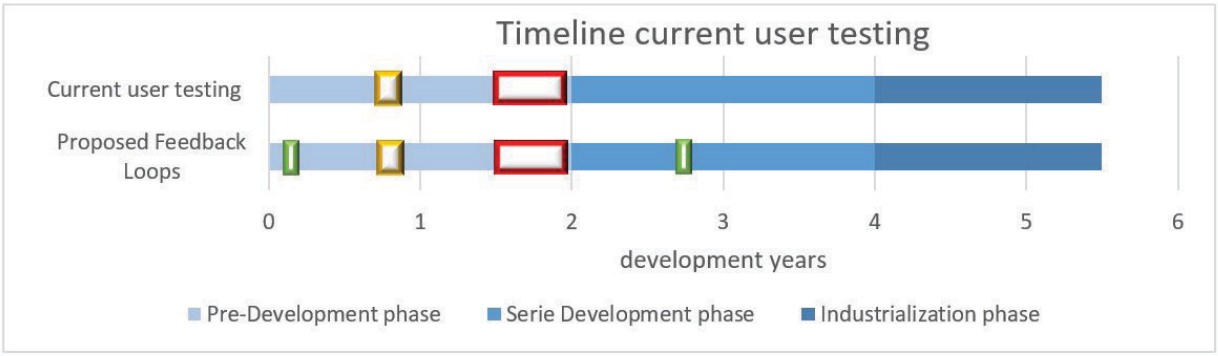


Figure 6 User test timeline

Data Analysis

One of the big databases that collect the majority of BMW's data is shown by a summary of the data available has been categorized. How and when the data is being gathered for this database resulted from an interview with an expert in charge of this database. What has been the conclusion is that the type of data is all very much quantitative and measurable. Still, there is room for qualitative creative data.

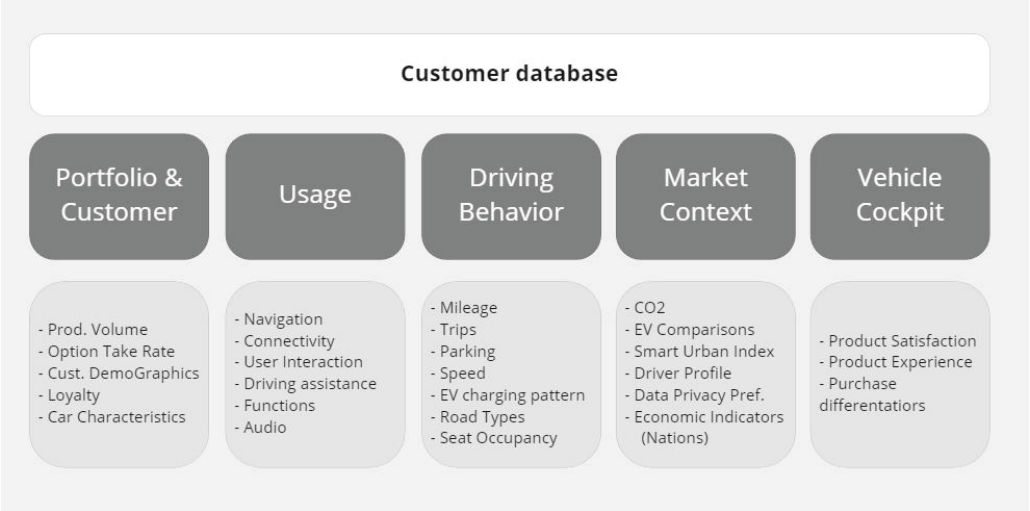


Figure 7 Database summary

Future Data Management

Next to the large database (figure 'Database summary'), "there are new developments in the global automotive industry that aim to extract interesting data"(Expert, 31). These are aimed at collecting data on customer experiences of market products and reflect what scores are being given on products or product features. These scores display what customers want, prefer, or talk about to their friends. BMW also recognizes this global trend and therefore, this shows that there are signs of interest in new types of data gathering. Types that can deliver more qualitative data to enable early insights for designers and developers.

3 - Problem Definition

BMW GROUP

From `Analysis 1´ conclusions come together and form a problem statement. As stated in the literature study, customer co-creation delivers creative feedback from the end-user and will be beneficial to the interaction between customer and product. From the company, analysis can be concluded that co-creation currently is very scarce in pre-development. Current user tests take a long time to prepare and even longer to analyze. On top of that, there is currently no method to gain customer feedback that stores the data in an easily implementable way. Methods to gather short-loop customer feedback to evaluate the input for early pre-development projects are therefore limited .

4 - Analysis II

4 - Analysis II

Trends and Opportunities

Trends and technologies that currently are promising are being closely monitored by a company like BMW. The most promising technological developments and most interesting trends from BMW research will be split into the `Tech. trends´ and the `Society trends´.

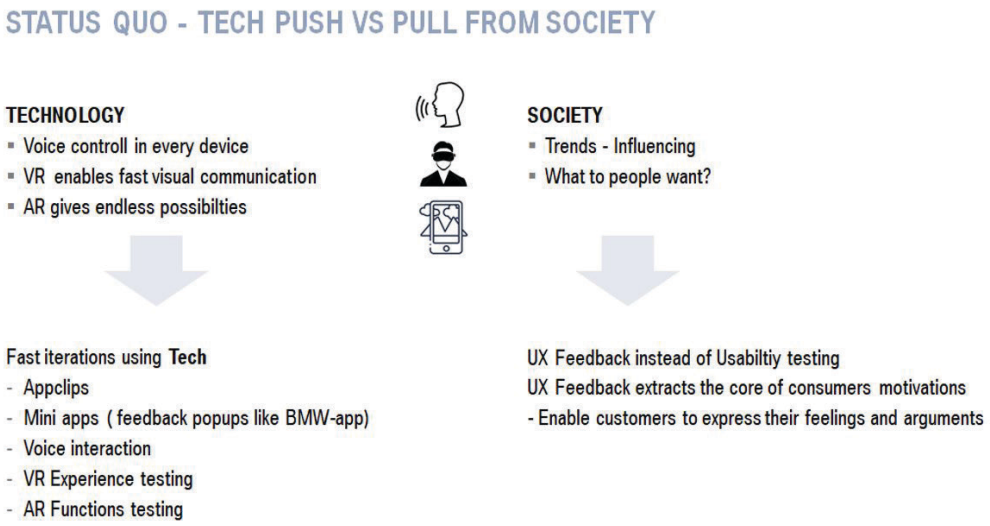


Figure 8 Technology and society

Society

The eagerness to be able to influence the world as an individual has been acknowledged by BMW and is being reflected in the brand identity. "Brands are created by customers, customers want brands that help them to fulfil their dreams, they don't want brands that sell them a dream." According to `The Brand and Customer Institute´ (BMW). To supply this demand, "BMW becomes a customer-centric brand." generally have a limited understanding of product innovations and are uncertain about the new benefits

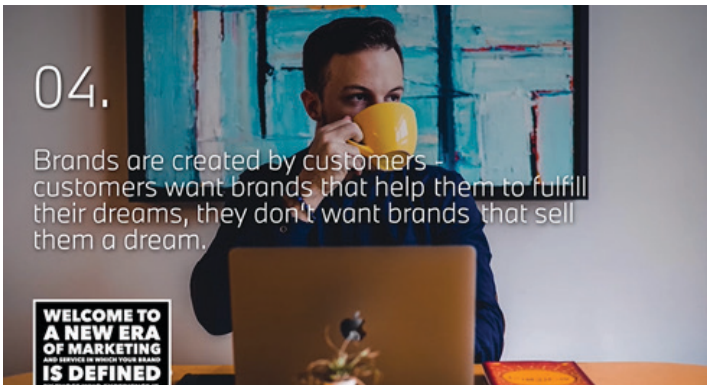


Figure 9 Society trends

The conclusion is that UX feedback will extract the core of customers' motivations. Customer feedback should enable the customer to express the core of their issues and enable them to explain where their personal experience originates from, not just enable them to make choices. By letting the customer first sum up positives, then negatives about a product. The customer needs to think of personal arguments in order, to sum up, their positives and negatives. By asking for solutions for their negatives this fits the trend of customers wanting to express their motivations and thus feel co-creation.

Technology

On the technology side, there are major developments seen in:

- Incorporating Voice User Interaction (VUI) in consumer devices
- Virtual Reality (VR) technologies enable fast visual communications
- Augmented Reality (AR) technologies enable endless visual possibilities
- Gesture User Interaction (GUI) enables users to interact with a product via (hand) gestures
- Haptic User Interaction (HUI) give vibrating communicative feedback to users

"Currently we see that all big software related companies dive into the world of Voice, VR and AR tools to develop new amazing UX products, of course, we follow these developments and need to keep up" (Expert, 2). As a result of this technology monitoring, many pre-development projects incorporate soft- and hardware tools that enable prototyping studies that make use of VUI and VR technologies.
Voice User Interaction (VUI) and Virtual Reality (VR) h ave been identified by the UI/UX departments as two key technology trends that are being explored more and more. Because automotive products are increasingly becoming more digital, a questionnaire cannot show the full potential of these digital products to customers. With novel technologies, the quality of customer feedback can potentially be higher and more interactive. In order to draw an unbiased conclusion at this stage, other technologies are also explored in this research. AR, GUI and HUI will be part of the status quo analysis and part of the resulting opportunity chapter in this research. In appendix 1.3 can be seen how the different technologies score.

Status Quo

To understand the contemporary situation, an analysis of the status quo will give a clear understanding of the current customer feedback interaction moments. The goal is to discover the currently available interaction points between customers and BMW. In the graph 'interaction moments' is an overview of the customer feedback interaction moments between BMW and their customers. Ranging from 'low fidelity' (basic functional prototype) to 'high fidelity' (product-like prototype) and 'rigid' (slow to adapt) to 'agile' (easy to adapt for a different purpose). The development stage where these interaction points take place is being displayed on the secondary horizontal axis, in order to understand gaps and possibilities.

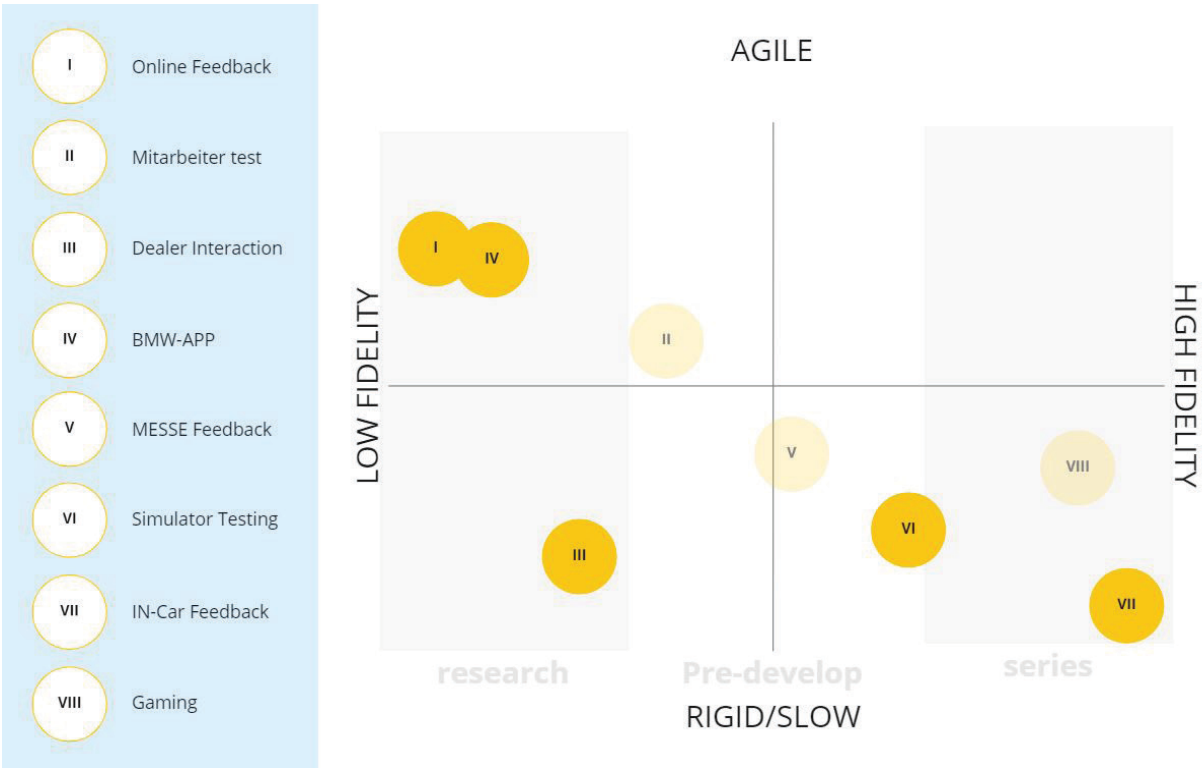


Figure 10 Interaction moments

Combine Interaction moments with Tech

In search of new opportunities, the current interaction types between customers and BMW are collected. Together with the knowledge gathered from interviews and information about department procedures and what could be improved, new possibilities were created and discussed with department employees. The current interaction types are each connected to new technologies to form a new interaction opportunity. Below in figure 12, the interaction types in yellow are combined with technologies in blue and form opportunity clusters 'I' until 'VIII'. For example, cluster 'V' contains customer interactions that are all tested on a fair (Messe), but by different technologies that enable these interactions.

	Interaction type					
Technologie		VUI	GUI	HUI	VR	AR
I	Online Questionaire	Voiceflow Questionaire	Gesture (Wii) Questionaire	Questionaire about haptics	Virtual Experience online	Augmented Experience online
II	Employee test	Pilot Voiceflow testing	Pilot Gesture testing	Pilot Haptics testing	Virtual Pilot testing	AR Pilot testing
III	Dealer Interaction	Dealer Voiceflow test	Dealer Gesture test	Dealer Haptics test	Dealer Virtual test	Dealer Augmented test
IV	BMW-APP	VUI Feedback menu in APP	VUI Feedback menu in APP	VUI Feedback menu in APP	VUI Feedback menu in APP	VUI Feedback menu in APP
V	MESSE-Feedback	Voiceflow tool on Messe	Gesture tool on Messe	Haptics tool on Messe	VR tool on Messe	AR tool on Messe
VI	Simulator Testing	Full Voice Interaction SIM	Full Gesture Interaction SIM	Full Voice Haptics Interaction SIM	Full VR Interaction SIM	Full AR Interaction SIM
VII	IN-Car Feedback	IN-CAR Voice feedback menu	IN-CAR Gesture feedback menu	IN-CAR Haptic feedback menu	IN-CAR VR feedback menu	IN-CAR AR feedback menu
VIII	Gaming	Voice assist gaming Simulation	Gesture assist gaming Simulation	Haptic assist gaming Simulation	VR assist gaming Simulation	AR assist gaming Simulation

Figure 11 Interaction types

Opportunities

In figure 'Shift in the position of interaction types' are possible changes illustrated in interaction types, enabled by new technologies. This chart is a result of a brainstorming session together with a BMW UI/UX employee. These changes illustrate new qualities enabled by combining a current interaction moment with new technology.

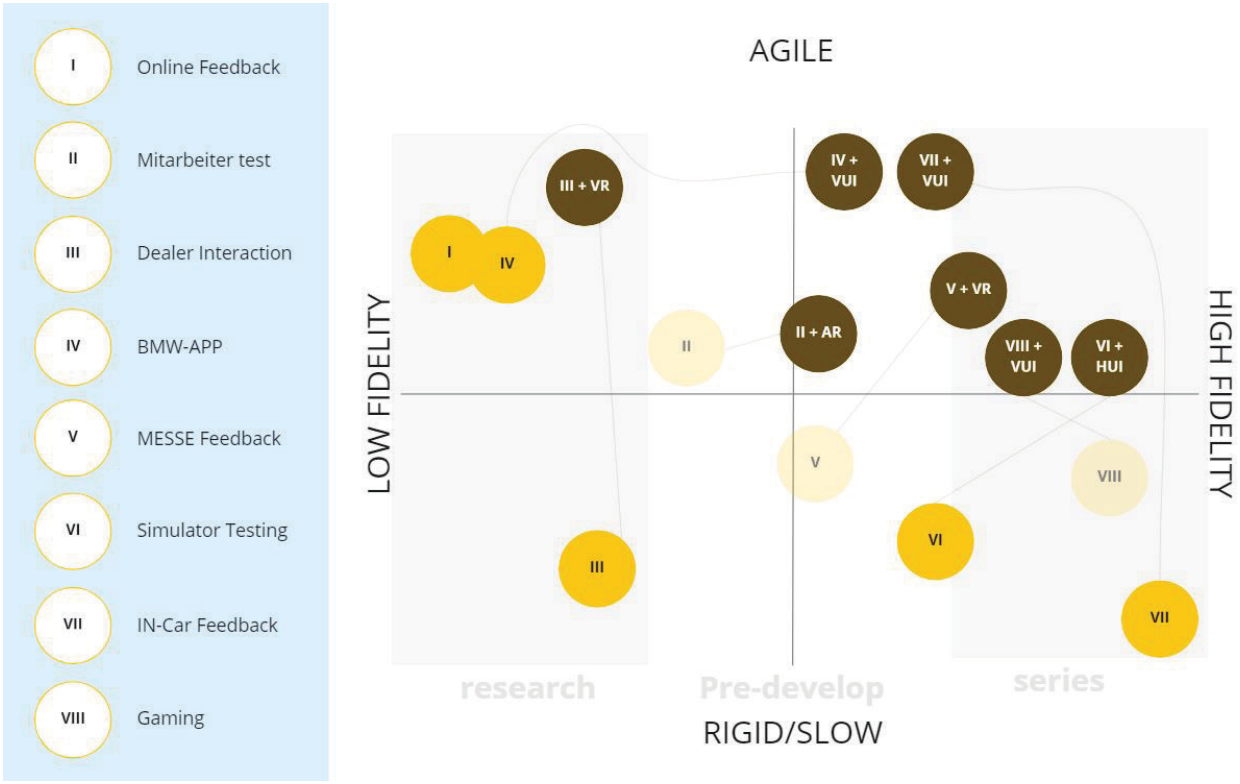


Figure 12 Shift in position of interaction types

For example, in the case of interaction number ´VII´ (In-Car Feedback), the change could be quite dramatic. The image shows a change from yellow to dark brown. Where dark brown shows a new location in the 3 axis diagram. This new location displays a possibility after new technologies are being implemented. In this case; currently, IN-Car Feedback would mean that there can be a development beta menu in the software of the car. This menu could display for instance a questionnaire or a layout where customers can type their feedback. This would take quite a lot of programming and time to adjust to different scenarios. On top of that, this process would need a lot of paperwork to get permission in order to be part of BMW´s identity. Therefore, located in the ´series´ section of the diagram and marked as ´slow´ and ´high fidelity. New prototyping software, could enable a drastic change of position in figure 12. If a feedback tool could use voice interaction software that is easy to implement, this will make this interaction moment shift to 'low fidelity' and enable 'pre-development' to benefit from this. Therefore, the location of an IN-Car feedback interaction type combined with agile voice interaction software could shift to the dark brown 'VII' dot.

Scoring the Opportunities

Following the brainstorming, a scoring process took place by categorization of the 40 new interaction opportunities. From this analysis, a conclusion could be made as to what can be the most suitable combination to target as a short-loop feedback opportunity. The table below shows a summary of all cluster groups that are being scored on six characteristics. Criteria that determined this analysis are:

1. Agility (easy implementable)
2. Exciting for Customer
3. Low Fidelity Iterations possible
4. Customer Bonding
5. Creative Data
6. Controlled Environment

		Agile (change Use Case)	Exciting for Customer	Low Fidelity Iterations possible	Customer Bonding	Creative Data	Controlled Environment	
I	Voiceflow Questionaire	6	6	7	6	6	4	35
I	Gesture (Wii) Questionaire	6	6	7	6	6	4	35
I	Haptical Questionaire	4	6	7	6	6	4	33
I	Virtual Experience online	6	7	7	6	6	4	36
I	Augmented Experience online	6	7	7	6	6	4	36
II	Pilot Voiceflow testing	7	6	9	7	7	7	43
II	Pilot Gesture testing	6	6	5	7	7	7	38
II	Pilot Haptics testing	5	6	7	7	7	7	39
II	Virtual Pilot testing	6	6	5	8	7	7	39
II	AR Pilot testing	6	6	5	8	7	7	39
III	Dealer Voiceflow test	8	8	9	8	7	8	48
III	Dealer Gesture test	7	9	5	8	7	8	44
III	Dealer Haptics test	6	8	7	8	7	8	44
III	Dealer Virtual test	7	9	5	9	7	8	45
III	Dealer Augmented test	7	9	5	9	7	8	45
IV	VUI Feedback menu in APP	8	7	9	7	7	6	44
IV	GUI Feedback menu in APP	6	8	8	7	7	6	42
IV	HUI Feedback menu in APP	7	7	8	7	7	6	42
IV	VR Feedback menu in APP	7	8	8	7	7	6	43
IV	AR Feedback menu in APP	7	8	8	7	7	6	43
V	Voiceflow tool on Messe	7	7	6	8	7	7	42
V	Gesture tool on Messe	6	8	6	9	7	7	43
V	Haptics tool on Messe	5	8	6	9	7	7	42
V	VR tool on Messe	6	9	6	9	7	7	44
V	AR tool on Messe	6	9	6	9	7	7	44
VI	Full Voice Interaction SIM	7	8	7	8	8	9	47
VI	Full Gesture Interaction SIM	6	9	6	9	8	9	47
VI	Full Haptics Interaction SIM	6	8	7	8	8	9	46
VI	Full VR Interaction SIM	6	9	6	9	8	9	47
VI	Full AR Interaction SIM	6	9	6	9	8	9	47
VII	IN-CAR Voice feedback menu	9	7	9	8	8	7	48
VII	IN-CAR Gesture feedback menu	8	7	5	8	8	7	43
VII	IN-CAR Haptic feedback menu	7	7	7	8	8	7	44
VII	IN-CAR VR feedback menu	8	8	5	8	8	7	44
VII	IN-CAR AR feedback menu	8	8	5	8	8	7	44
VIII	Voice assist gaming Simulation	8	8	5	7	8	7	43
VIII	Gesture assist gaming Simulation	7	8	5	7	8	7	42
VIII	Haptic assist gaming Simulation	5	8	5	7	8	7	40
VIII	VR assist gaming Simulation	7	8	5	7	8	7	42
VIII	AR assist gaming Simulation	7	8	5	7	8	7	42

Table 2 Scoring opportunities

From analyzing the results, Dealer Voiceflow tests and IN-Car Voice Feedback are the highest scoring combinations. Therefore, these were discussed with BMW UI/UX employees. The BMW Covid policy on the global Covid situation, unfortunately, has ruled out Dealership tests as a research possibility. Considering the fact that voice technology will be less influenced by this situation, BMW employees considered this as the most interesting cluster to proceed with this research in. On top of that, as seen from the data; all opportunities combined with voice technology scored relatively high. Therefore, the decision has been made to continue the research on voice interaction (VUI) as a technology enabler.

IN-Car VUI feedback

After analyzing what the most concrete opportunities could be. Further decision-making has been done to choose a customer feedback opportunity: IN-Car VUI feedback. This customer feedback method will use `Voiceflow´ software running on Alexa and can possibly be implemented inside a series-production BMW car, via contemporary internet connections, called `BMW Connected Drive´.

5 - Concept Phase

5 - Concept Phase

BMW GROUP

What to Verify

The developed concept should be able to verify the following research questions.

RQ1: More creative input can be expected when using methods with higher levels of Natural User Interactions* in customer feedback tests.

H1 Hypothesis: Expected is that a method that makes use of higher levels of Natural User Interaction* in customer feedback, will result in customers giving more creative feedback.

RQ2: Can customers experience a higher value of co-creation by methods with higher levels of Natural User Interactions in customer feedback tests?

H2 Hypothesis: Expected is that a method that makes use of higher levels of Natural User Interaction in customer feedback, will result in customers experiencing a higher level of co-creation.

RQ3: Will the implementation and the analysis of customer feedback be less time-consuming than current user tests when using IN-Car VUI Feedback as a method?

Introduction of Three-method feedback test

Resulting from the past research, a three-method customer feedback test has been created to explore the most successful implementation. The first method will be a customer test that makes use of a contemporary feedback method that is my research, represents the main method that is being used by BMW, a questionnaire. The second step will differentiate from the first one by using voice interaction to facilitate the customer feedback session, in order to grow towards a more real-world user interaction. Between these two steps, differences in the outcome of these customer feedback sessions are being tested. The third step is to create the ultimate Natural User Interaction by implementing the voice interaction feedback session inside a series-production BMW. Therefore, the feedback about car functions is being generated in an actual car and therefore in the most realistic environment possible. Below is an illustration of the three method feedback test.

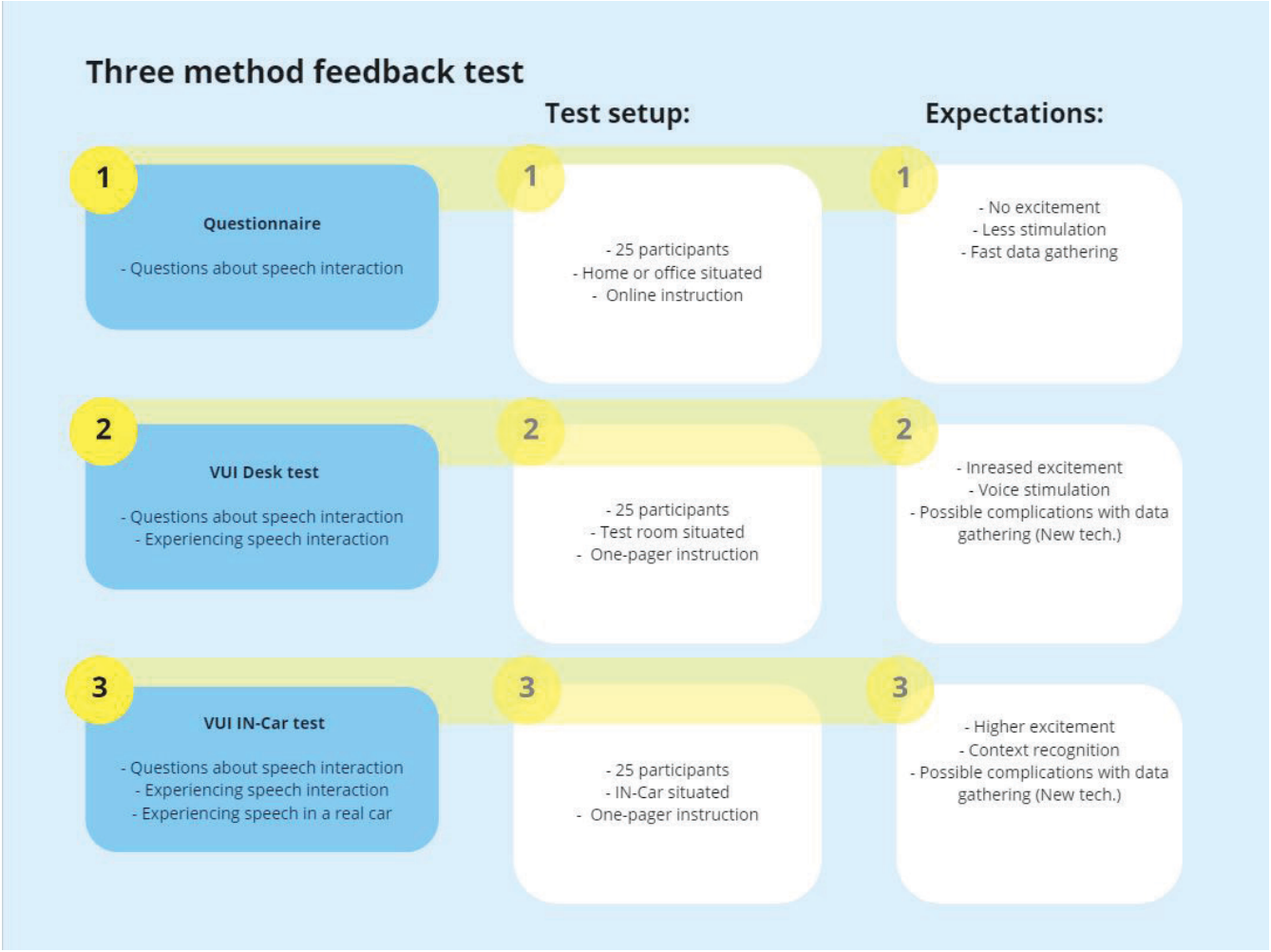


Figure 13 Three method test

This figure shows that the next customer test gets a more realistic scenario by actually interacting via speech and in the final method actually experiencing speech in a car. Each step contributes to a better Natural User Interaction. The questionnaire will require the participant to imagine the scenario of being inside a car and using speech interaction. In the VUI desk test, the participant actually interacts with the speech interaction by answering questions about the product speech interaction. The IN-Car VUI test is giving the most input to the participant by offering a real-world scenario by actually using speech interaction inside a BMW. Differences in creative feedback will showcase if higher Natural Interaction will change the creative output participants deliver in these customer feedback customer tests.

Customer test implementation

To facilitate the customers' test during the current Covid pandemic, the aim is to test as much as possible online. Since it remains unclear how long a (covid) pandemic will be among us, future implementation of gathering customer feedback online, instead of face-to-face held questionnaires becomes more and more appealing. Next to this, there is a high demand for speeding up the process of gathering and analysing customer feedback inside BMW. Therefore, aiming for online and more automized solutions is desirable.

Participants

The 3-method test is set up with n= 25 participants per group, which makes n= 75 participants in total. This amount is chosen in order to be able to analyze the data qualitatively, while simultaneously ensuring to have a large enough number of participants to validate possible drawn conclusions. All participants were native Germans and therefore all three customer tests were translated into German. The participants are supplied by an internal BMW customer testing department, where presets in the search for employees can be selected. For this research, participants could not have a relation to the development of innovations regarding customer interaction products or features.

is giving the most input to the participant by offering a real-world scenario by actually using speech interaction inside a BMW. Differences in creative feedback will showcase if higher Natural Interaction will change the creative output participants deliver in these customer feedback customer tests.

Setup

To let the actually used method be the main variable between the three customer tests, the content needs to be similar and therefore the content of the questions are the same. On top of that, the participants need to be prepared and informed in similar ways. Each introduction contained that the subject of the customer test was going to be "Sprache-Interaktionskonzept " (speech interaction concept). Since not every participant would be familiar with speech interactions, a short explanation of this technology was given to them; „Untersuchungsthema ist der Sprachassistent des Autos. Diese Technologie erlaubt es, Infotainment-und Fahrzeugfunktionen durch Sprachbefehle zu bedienen". The introductions (shown later on) to the test, therefore, have been the same, apart from the fact that some practical instructions for how to interact with the speech interactions. Specific instructions were necessary because the speech interactions were created especially for this research and therefore were pre-development prototypes that needed some instructions in order to enable smooth operations. Following, the differences are discussed separately.

Questionnaire

The decision to test the questionnaire as a benchmark for the two speech interaction tests is because the questionnaire is currently mainly used to gather information from customers inside BMW. This is sometimes the primary source of information in customer tests and on other occasions a source of feedback to enable the participant of a customer test to evaluate a product experience. The repeating factor is this questionnaire method to gather information for BMW, resulting in the decision to use a questionnaire as the contemporary method against the two novel speech interaction tests.

The pre-selected participants need to follow up a few instructions for the questionnaire: "Bitte finden Sie einen Platz, an dem Sie mit WLAN-Verbindung und 20 ungestörten Minuten sitzen können." (find a seating place, Wi-Fi connection, ten minutes of uninterrupted time). This instruction was given to simulate a more controlled environment, where the participants could concentrate on the questionnaire. The question remains whether every participant was able to follow up on this, regarding the current home-office situations, but this is had to be accepted for this research

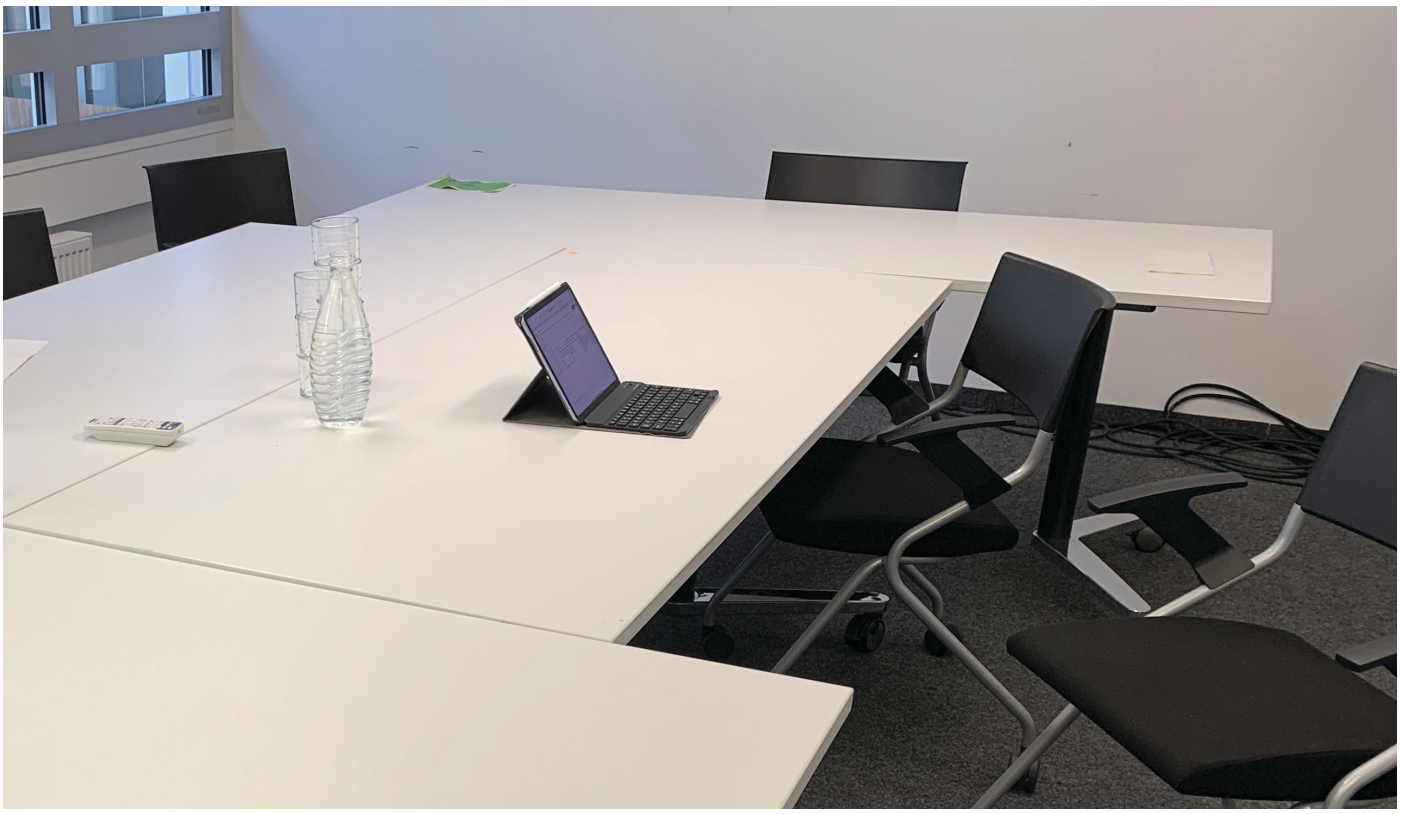


Figure 14 Questionnaire setup example

VUI Desk Method

The VUI Desk method test is located in a BMW office, this could not be facilitated online since an Amazon Echo device was needed to run the design speech interaction. Therefore, the participants were invited to a meeting room in our department. The Additional needed tools are Audio hardware; microphone and speakers (Amazon Echo), internet connection, Voiceflow software and an Amazon Alexa Developer account. The figure 'VUI Desk setup' shows this setup of a person sitting at a desk in a closed BMW office. The person interacts with the Alexa skill via the Amazon Echo hardware.

Specific to this customer test is that the participants received instructions to when they were able to answer, there was a given answering window. Participants had to wait for the blue bar to appear at the bottom of the Echo screen, this indicates that the answering window is opened. Figure 'VUI IN-Car introduction' shows this instruction one-pager.

Currently, the participants are not able to time when they start their answer themselves, this is a limitation of the Alexa developer software. In the `Pilot testing´ chapter under `Thinking seconds´ will be explained how there has been created a workaround for this limitation.

Figure 15 VUI Desk setup

VUI IN-Car Method

This third test will take place in a stationary series-production BMW with BMW Alexa software running on it. Next to the BMW car, extra needed tools are necessary to run the Alexa skill on the BMW: A mobile phone with a `My BMW´ app that is linked to an Alexa Developer account. In short: This mobile phone is connected to the BMW car over Bluetooth and runs the Alexa skill over the internet network of the phone. In this way the participant interacts fully realistic over the hardware of the BMW car, no "Wizard of Oz" tricks are being used during these tests.

Specific for this test is that the speech indication for the speech answering window is not a blue bar, but a microphone indication in the head-up display of the car. In this way, the participant is indicated when to start answering. This, therefore, was explained before the test started by a short introduction one-pager shown in the figure 'VUI IN-Car introduction'.

Apart from this difference the VUI content is unchanged, the same Alexa skill runs on both the Amazon Echo device as in the BMW car. A look into the setup of the third customer feedback test method is given by the 'VUI IN-Car setup' figure.

Figure 16 VUI IN-Car setup

VUI Customer test

The chosen Voiceflow software enables programmers to design a speech interaction that can interact with a customer and ask questions, the customer can answer, via a hardware tool, by giving spoken feedback. The designed research set-up starts with easy 'yes/no' questions to let the participant familiarize themselves with how the VUI functions (Carbon, 2005). Then follow some choices that decide on what voice-guided feature the customer has the most experience with. The program reacts to this answer and this answer influences the next steps in the flow. In this way, the participant can determine their own conversation flow in the VUI. To simulate a Natural User Interaction, the conversation could ask personal experiences of the person and then is able to use this information and apply this at a later moment in the conversation.
From a research perspective, the content parts can be marked for different purposes. The next figure 'Test content' illustrates the different purposes of each stage of the content of the customer tests: `Inform´, `Data´, `Familiarize´, `Feedback´and `Narrative´.

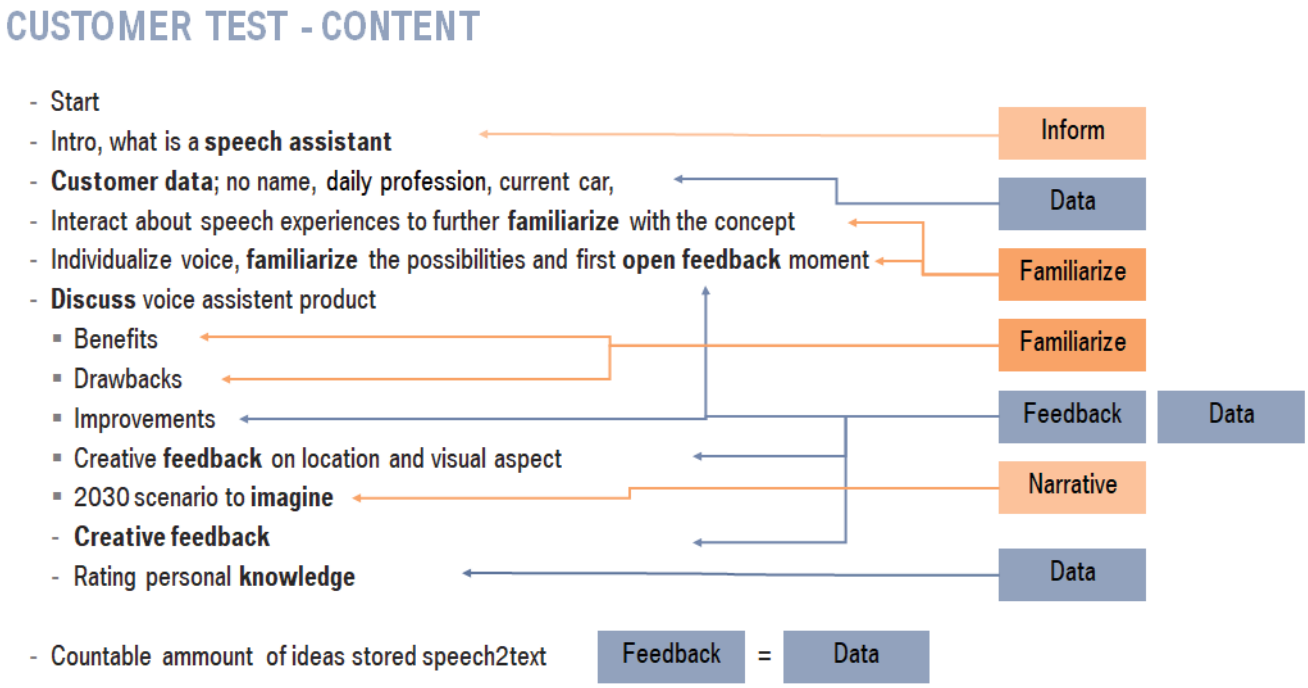


Figure 17 VUI content

On the left side, the content of the interaction flow starts, parallel with the purposes on the right-hand side, from top to bottom. The introduction informs the participant about the topic `speech interaction´ and describes the meaning. Soon after the introduction, basic personal data is being gathered and shortly after these data questions, the familiarization with the possibilities of speech interaction is started. An example question is: „Haben Sie bereits ein Fahrzeug mittels Sprachinteraktion bedient?" (Have you ever used an car feature activated by speech interaction?) with an follow-up question: „Zu welchem Zweck verwenden Sie die Sprachinteraktion hauptsächlich? Bitte wählen Sie Zwischen "Ziel verwenden", "Radio", oder "Telefonfunktionen" (With what purpose do you use speech interaction mainly? Please choose between "setting a destination", "Radio" or "Phone functions"). The questions are not very important for this research but enable the participant to think about their previous experiences and further possibilities. Therefore starts off the familiarization phase.

Secondly, the participant is able to alter the sound of the voice to further experience possibilities of speech interaction. Then the participant is challenged to think of as many benefits as possible a speech interaction can provide in a car, followed by as many negatives as possible. And thereafter the participant is challenged to make use of their creativity in order to think of solutions that could solve their mentioned negative aspects. In this research this is referred to as the `first creativity question´ and the first creative feedback the participant gives that will be analyzed and compared to the outcome in the other methods. Thereafter, user input for the UI/UX department is implemented by asking questions about the visual aspects of speech interactions. This is not part of the creativity research but contains nevertheless interesting data for developers in the department. Following this, there will be a short time that the participant is asked to think about a future 2030 scenario where their car will be able to assist the participant in their own specific life. This enables narrative transportation by letting the participant think deeply and visualize their life and own individual benefits by car functions activated by speech interaction. This is the last creative data gathering point that will be assessed in this research, later on, referred to as the `second creativity question´. Finally, there is more personal data gathered by asking to what extent, after taking part in this study, the participant thinks they are able to influence BMW product development. This is being asked at the beginning and at the end of each customer test. In other to monitor growth in this grade from before and after the study, within the same methods. On to op that to measure differences in grading the ability to influence product development between the three methods.

Method testing

To test the VUI test as a method, this prototype tests if creative data can be stored in speech2text making use of the prototype. This prototype is designed on a software named `Voiceflow´, this software runs over an Alexa developers account and stores the spoken feedback in a Google sheet. And this has been tested successful, spoken text is being stored in desired rows and columns in a linked Google sheet file. Therefore, we can store speech2text customer data directly, the programmed feature is shown by the screenshot below in figure 18 'Speech2text storage in Google sheet'.

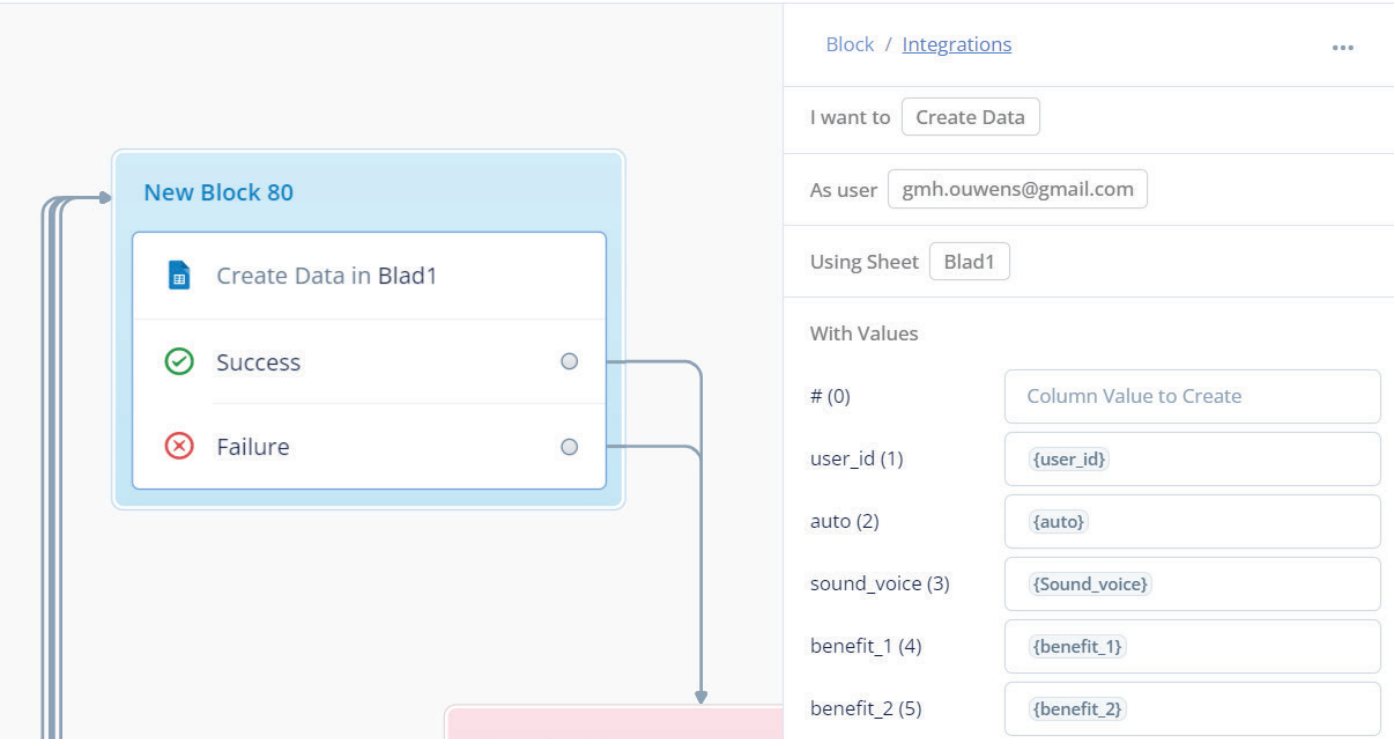


Figure 18 Speech2text storage in Google sheet

After the first successful tests, the potential of exploiting speech technology for customer feedback became clear. However the developed Alexa skill is still a prototype, but the benefits are crystal clear. The Alexa skill currently is able to store speech data and directly transform this into text data in a controlled manner. Therefore, the data can be understandably presented to developers or analysts. "The time-saving element will deliver great benefits. Since, previous early user test programs often failed due to high labour intensity of executing the feedback sessions and most of all processing of the gained data" (Expert, 14). In the `Results´ chapter of this research more elaborate analysis of the time-saving potential will be discussed.
Hand in hand with time-saving goes a money-saving aspect. "If automated customer feedback storage will be implementable in the future this will save a lot of manhours" (Expert, 14) and therefore result in research money savings. However, for this research, the money aspect will not be discussed elaborately.

Product testing

While testing the method, VUI enables to test of the product at the same time, or in the same customer test. The prototype tests this by asking the consumer to fine-tune the sound of the programmed voice. The possibilities are vast. The customer test is enabling the user to adjust the sound of the voice in tone between manly and female and a more soft tone and between a fast or slow speaking pace. These are examples I use to let the participants experience opportunities and therefore stimulate their brains about the possibilities of speech interactions. However, many more options can be programmed.

Pilot testing

To iterate on developing the execution of the thee methods of customer testing, a pilot testing phase by exposing my customer tests to internal BMW UI/UX experts was the first test with employees from outside the department where the tests are designed. This resulted in knowledge that indicated current shortcomings, future opportunities and improvements that could be fastly implemented.

#	**Expert knowledge**	**Contribution**
1	Product Owner UI Vision	Time management of answering and future vision
2	User test expert	Future implementations and distractions and German language
3	UI entertainment	Developer benefits if data is exactly relatable to user issue input
4	UI system PO	Finetuning of questioning
5	Late user testing expert	Language recommendations and participant selection
6	UI Passenger entertainment	Time pressure indication, overall setup participant tests
7	Alexa programmer	Technological back-end programming solutions
8	UI Concept Designer	Question content and future implementations
9	UI Basic Concepts	Future implementations and research possibilities
10	Alexa developer	Speech recognition failure solving

Above is a summary of what kind of experts took part in the pilot tests and what they contributed. The participants took part in pilot tests as a participant would do. The introduction one-pagers were used for informing purposes, followed by the actual test. This step enabled further development of the Alexa skill since the experts foresaw both technical difficulties and gave tips on changing the questioning.
The knowledge gained by the pilot tests resulted in multiple improvements, problems to solve and future implementation advice in my rollout strategy. Following the points of improvement are being explained.

Thinking Seconds

As the first few test pilot tests were executed, the first attention point was clear. Participants needed more thinking time. The biggest limitation of the Alexa software is the fact that the answering window is only eight seconds and can not be altered. To create solutions, or maybe even call it a workaround, bigger pauses between question parts were programmed to enable the participant more time to answer the questions. For the most complicated questions, there is a thinking pause programmed for eight seconds straight. After finetuning the pause moments there is now a smooth conversation possible. The 'Thinking time' figure below illustrates how the implementation of pauses between sentences and the beginning of an answering window is applied in the software by the 'break time seconds'.

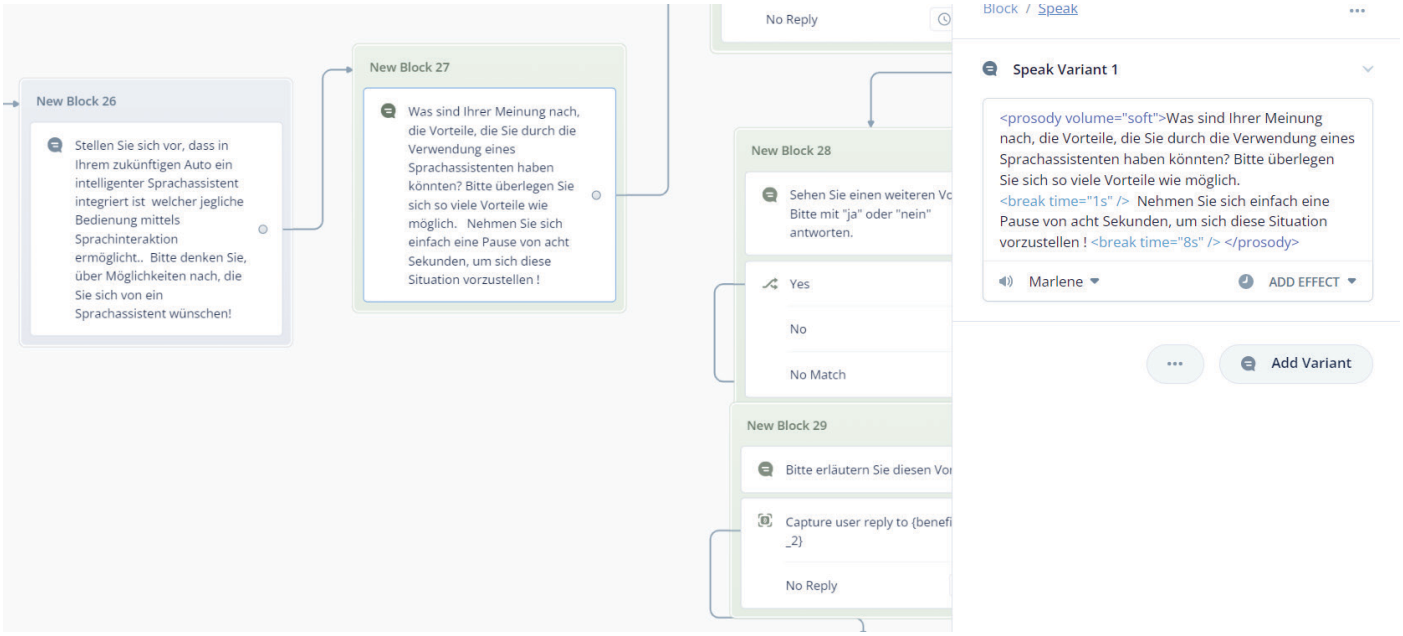


Figure 19 Thinking time

Multiple Reaction Promts

To prevent the conversation from getting boring, programming different variations on the question replies was needed. A human also likes to vary in word choice if multiple times the same reactions are possible. Programming multiple reactions prompts makes the conversation more interesting.

Connection Steps

Compared to a questionnaire we use more words in a conversation. This is because in a questionnaire it does not seem strange if there are questions asked without too much connection or introduction words. The experts noticed there were more connection steps needed between the questions in the VUI tests. Therefore, clarifying connection words or sentences were added. As shown by the screenshot below by 'New Block 19', this sentence would not be needed in a questionnaire but has a great clarifying purpose in a speech interaction.

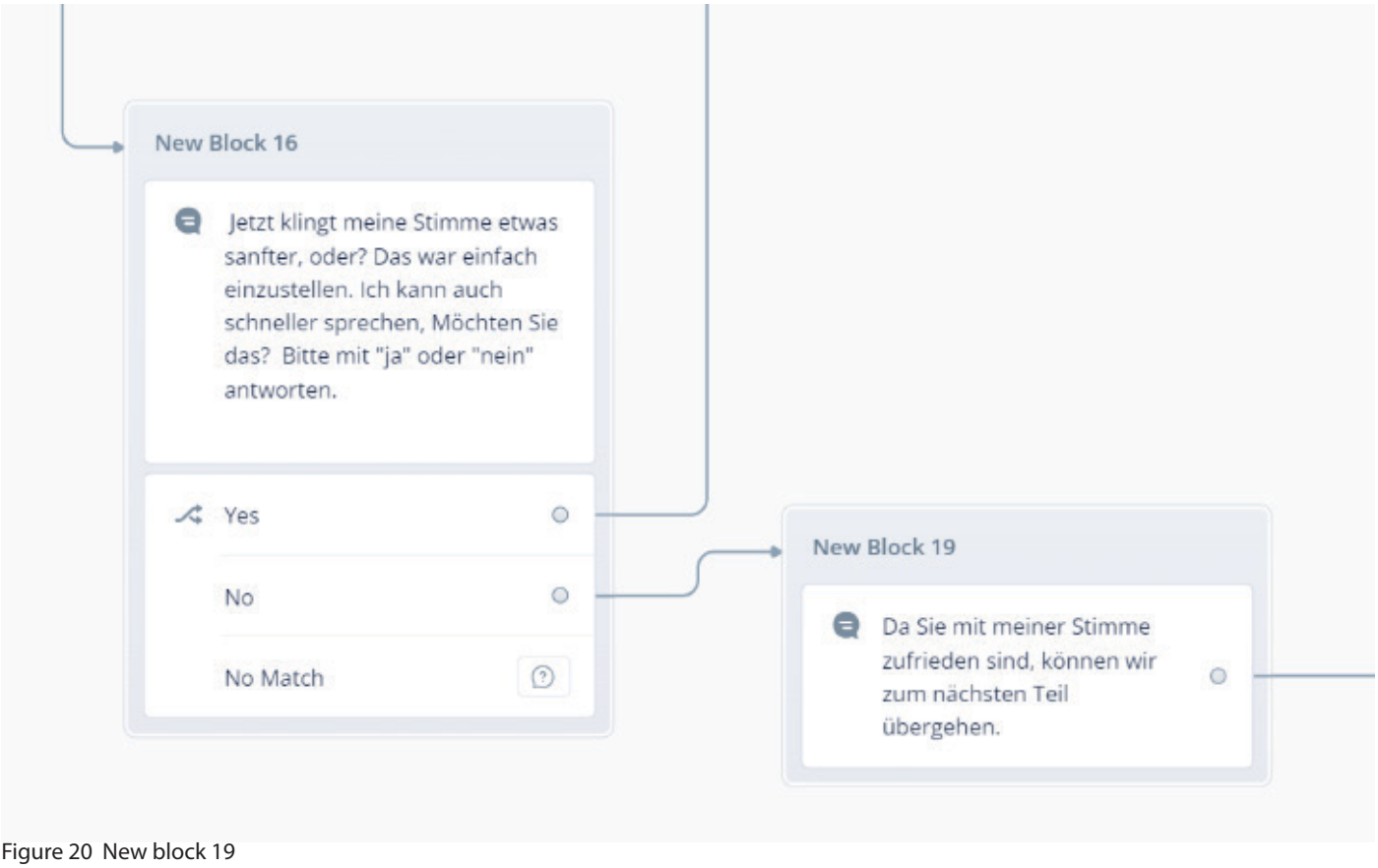


Figure 20 New block 19

Utterances

Utterances are programming options that the prototype can recognize, these words function as triggers to go to the next action. Below is shown that if the customer says “calling someone”, “to call someone”, “to call”, “mobile phone” or just “calling” the prototype is triggered and selects the option capture_phone_ functions’. Therefore, the flow follows the line from ‘capture_phone_functions’ to the next block. Adding more utterances increases the chances that the speech interaction will understand the input of the participant and continue the conversation. Therefore, multiple iteration tests resulted in improvement on this subject. A screenshot in the figure ‘Utterances’ gives an impression of the software function.

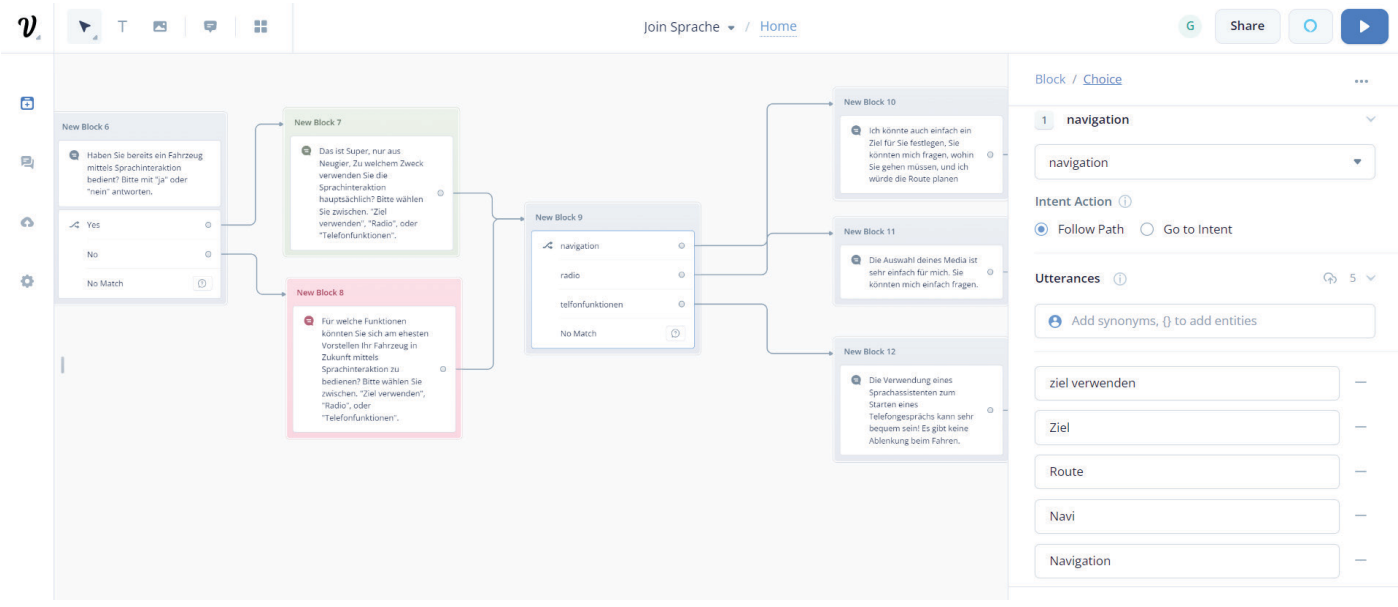


Figure 21 Utterances

Conclusion

The first conclusions are that this is a realistic tool. On top of that, this will be an interesting aim for BMW: A method that enables developers in pre-development stages to get direct feedback from customers or internal employees, including automatic speech to text storage. This direction included programming the whole skill, and these have been applied and are realistic to be cunsulted and function for tests with employees.
The questionnaire test will be held online and the VUI tests run over Alexa Echo hardware and respectively Alexa over BMW Connected Drive. In this way, all the planned aspects can be tested.

6 - Evaluate

BMW GROUP

Expected Results

By evolving the amount of Natural User Interaction* of each customer test towards a more real-world experience, it was to be expected to influence the amount of creative feedback positively (RQ1). To validate this as accurate as possible. Eliminating other parameters has been a goal to achieve as much as possible. A questionnaire is common to be self time-paced for a customer, and the time in an automated voice interaction is machine-paced. "If customers would feel time pressure this will influence the output they give" (Expert, 4). Therefore, iterations of the VUI setup are made to limit this time pressure aspect. Enabling the customer to make use of thinking seconds reduces this problem and becomes the machine paced interaction partly self time-paced.

"It is to be expected that customers feel that their opinion is being taken into consideration and experience this positively" (Expert, 8). By enabling the participant to score the influence they think they have on the product development of BMW before and after the customer tests, this grasps if these cases would influence the perception of influencing BMW product development. This forms the answer to RQ2.

The customer feedback model will be expected to deliver validations to pre-development projects, this will supply the demand for early validations. Therefore, it is to be expected that this tool is less time-consuming than current user tests. A timespan of one to two months would be a breakthrough. Comparing the time steps in the research methodology in the current user studies with the setup needed to implement VUI as a customer feedback tool will answer this question.

RQ3 will be answered by findings from the expert interviews, where useful insights in timespending related costs and implementation procedures were discussed. However a great art of this information is considered to be sensitive and for instance the costs will not be part of this report.

Summary

What are the circumstances and assumptions that needed to be taken into consideration? A task was to understand to what extent the change of environment; from a home desk while filling in a questionnaire or interacting with a VUI or to a BMW car, does influence my conclusions. The conclusions will be made on the effect of steps in applying an environment on the expected increase of creative feedback from the customer.

7 - Results

7 - Results

Research Question 1

The results of this research can be divided into three different parts that each answer, one research question. RQ1 is being analysed by two approaches: A design-oriented analysis and a marketing-oriented analysis.

Design analysis

The design-oriented analysis is being executed by grading the feedback of the participants on three criteria; fluency, flexibility and originality. "These are three of the four basic elements of divergent thinking"(Gonçalves et all., 2013).

Fluency is analyzed by defining the number of ideas produced by the participant per question in the tests (Guilford, 1950). Therefore, the number of ideas will be analyzed per participant, if the participant is able to generate a higher amount of ideas this is regarded to be more valuable feedback according to this analysis.

Flexibility is the capacity of switching between domains and therefore being able to grasp different approaches to a problem solution (Guilford, 1950). For this reason, the ideas are analyzed by distinguishing if the participant ideas belong in different categories, to determine if participants create ideas from multiple domains.

Originality is defined by Guilford (1950) as the ability to create novel and uncommon ideas. Amabile (1996) describes that originality is an important aspect to define creativity, together with the appropriateness and usefulness of the idea.
This originality score is calculated by the average occurrence of the ideas of a participant compared to the other generated ideas. Divided by the total amount of generated ideas (50).

In the design analysis table the discussed criteria are explained by examples. The Fluency criteria show that person `1´ generated two ideas and therefore scores `2´. Secondly, in the Flexibility criteria person `1´ generated ideas that are ranked as very similar and therefore both belong to category `A´, this results in a score of `1´. Finally, person `1´, scores in Originality by generating ideas number 30 and 23 of the total amount of 50 independently generated ideas. Resulting in the calculation that gives a score of 0.53 ((30+23)/2/50= 0.53).
The grading and counting of the generated ideas were being executed while being blind to the conditions. Since this contains no in-depth qualitative analysis of the feedback it has been decided that this was acceptable. The `Design analysis´ enables this research to compare the outcome to the `Innovation management analysis´ and increase the amount of gathered data. On top of that, It seemed interesting to compare the outcome of the two methods and see if they generate comparable significant results.

Fluency

Person	idea	idea	idea	idea	idea	Fluency
1	I	II				2
2	I					1
3	I	II	III	IV		4
4	I					1
5	I	II	III			3

Flexibility

Person	idea	idea	idea	idea	idea	Flexibility
1	A	A				1
2	B					1
3	C	A	A	C		2
4	B					1
5	A	C	B			3

Originality

Person	idea	idea	idea	idea	idea	Originality
1	30	23				0.53
2	28					0.56
3	7	25	30	22		0.42
4	10					0.2
5	17	28	13			0.39

Table 3 Design criteria

Innovation Management Analysis

According to innovation management literature, the key qualities to measure creativity are novelty, customer benefit and feasibility. ´Novelty´ of the idea reflected against contemporary existing products, the ability to solve the underlying problem forms the ´Customer benefit´ and the possibility of translating an idea into an actual product is called `Feasibility´. (Poetz & Schreier, 2012).

Experts grading creativity

The generated data using the innovation management criteria have been rated by two independent BMW experts just like the approach of Poetz & Schreier (2012). In this way, the gained data results have been assessed by UI/UX authorities. The experts rated the criteria on a 1-5 scale. Where 1= ´low novelty´/´low customer benefit´/ ´low feasibility´ and 5= ´high novelty´/ ´high customer benefit´/ ´high feasibility´. Also according to the approach of Poetz & Schreier (2012), the experts could label an idea with a value ´0` when they assessed the input not to be a true idea (but more as input on the topic)..

	Ideas to solve speech function problems	Novelty	Customer benefit	Feasibility
Person	Idea 1	1: Not Novel - 5: Very Novel	1: No Benefit - 5: High Benefit	1: Not Feasible - 5: Very Feasible
1	Die Möglichkeit an verfügbaren Befehlen erhöhen	2	3	3
2	Mikrofon auch auf Beifahrer ausrichten gute Anbindung ans Internet und ggf. lokale Speicherung von gewissen Daten	3	3	4
3	Algos, die Störgeräusche reduzieren. Training auf eigene Stimme. Alternative Bedienoptionen zur Verfügung stellen.	3	4	3
4	So viele Begriffe wie möglich in die Programmierung der Sprachsteuerung einfließen lassen.	2	3	4
etc..				

Table 4 Expert scoring example

The experts are being introduced to these criteria by a short instruction manual, which can be found in appendix 1.5. In this instruction, the innovation management theory and criteria, novelty, customer benefit and feasibility were explained. Secondly, the experts were instructed to practice the rating by completing an exercise of rating ten generated ideas from the pre-test. From that, they were asked to discuss their approach with each other and come to comparable rating tactics. Finally, the experts were instructed to rate all the generated customer feedback individually. Therefore, this resulted in two dataset ratings of the same content. The experts were unaware of the fact that the data has been generated by three different studies and consisted of two creative questions. The first creativity question asked the participant to think of solutions for their given negatives on speech interaction in cars and the second creativity question asked them to think of future features fitting their personal life. The data has been randomized and therefore experts could only rate the data on the idea itself. A more elaborate explanation of the analysis of the dataset follows in the next chapter.

Data Analysis

The expert datasets needed to be analysed together and therefore conveniently merged in one datasheet. In order to justify this step, a correlation test between these two datasets was needed to be performed.

Pearson Correlation test

The correlation test makes sure that the created datasets of the individual experts do not show significant differences. The Pearson Correlation test of the data should at least give an r-value between 0.5 and 0.6 and have a p-value of 0.05 in order to be significant. On all three criteria, the datasets of Expert 1 and Expert 2 approached the significance. The results of; Novelty ($r = 0.67$, $p < 0.001$), Customer Benefit ($r = 0.57$, $p < 0.001$), Feasibility ($r = 0.58$, $p < 0.001$), show that the requirements are being reached and therefore I assume that the datasets are justifiably comparable to merge them together

and continue the research with the average values between Expert 1 and Expert 2. In this approach every created idea has an average score on each criterion (novelty, customer benefit and feasibility), this average score is being multiplied to enable a score per idea (avg (N*CB*F)).
Participants were able to generate multiple ideas, therefore all ideas were part of the data. Finally, a summary of the idea scores per person is divided by the number of ideas generated by that particular person. This was executed according to the innovation management criteria (Poetz & Schreier, 2014) and resulted in the average score per idea per person (VGM AVG value).
Outlier analysis using z-scores, where -3<z<3. Z-scores check showed me that there was only one critical score Z = 3.14. This single high score could be clarified by a relatively high score (4 out of 5) on each innovation management criteria and therefore achieved a high average score. For this reason, it was regarded as part of the research and remained this high score in the dataset, it was only slightly deviating from 3.

One-way ANOVA

One-way ANOVA tested the differences between methods to explore the effect of the method on creativity scores (RQ1). Levene´s test of homogeneity of variances showed no significance (p > .05), allowing for the one-way Anova test. The data upon which this conclusion has been drawn upon is shown in appendices 1.6.2 and 1.6.3.

First creativity question

For the first creativity question, one-way Anova showed a significant difference between the methods on the Fluency (F(2,53) = 3.11, p = .05) and Flexibility (F(2,53) = 3.54, p < .04) criteria. Originality criteria (p > .38) and the innovantion management creativity score (p > .16) showed no significant differences between the methods.

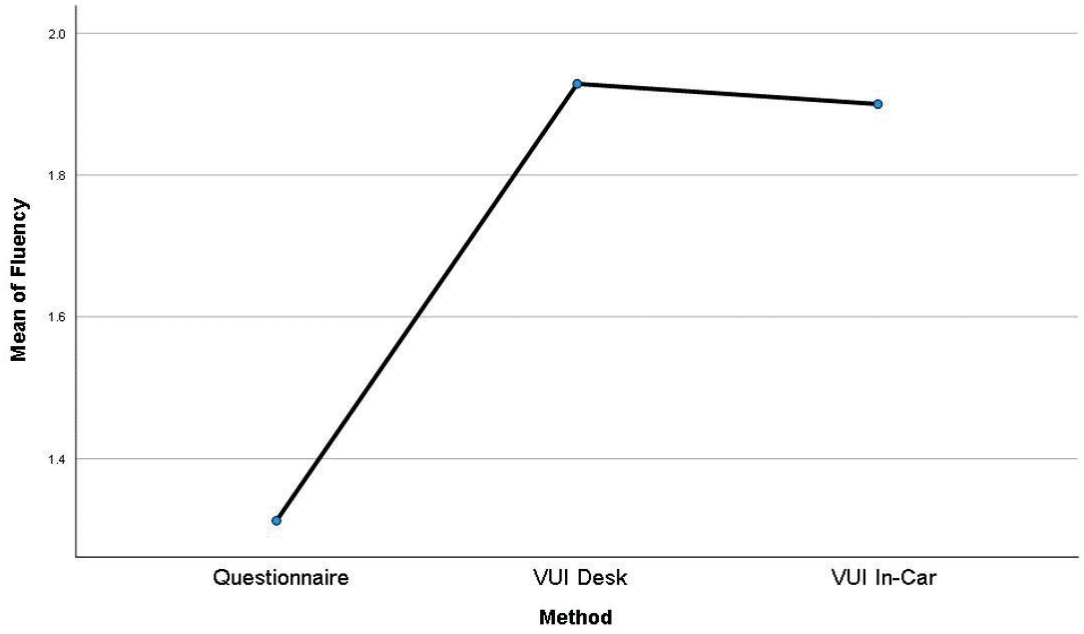


Figure 22 Fluency

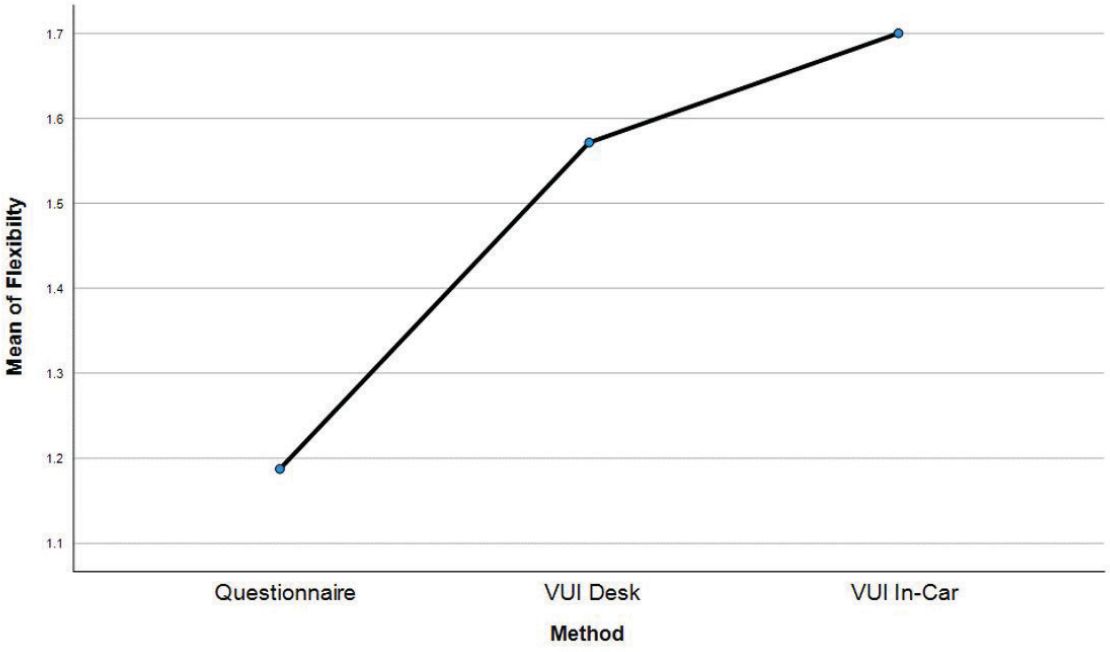


Figure 23 Flexibility

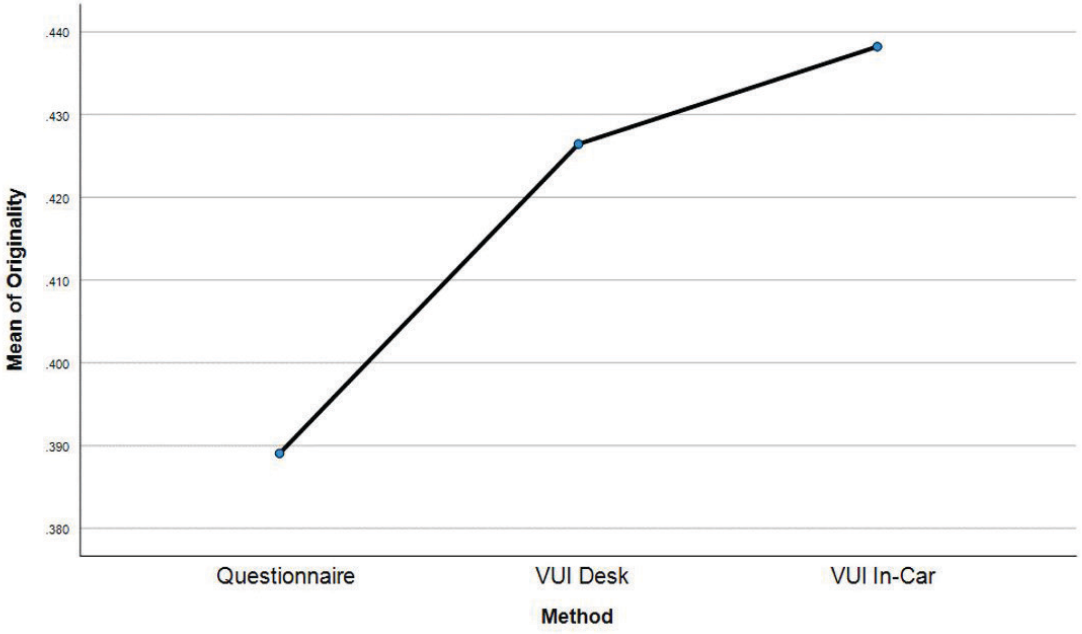


Figure 24 Originality

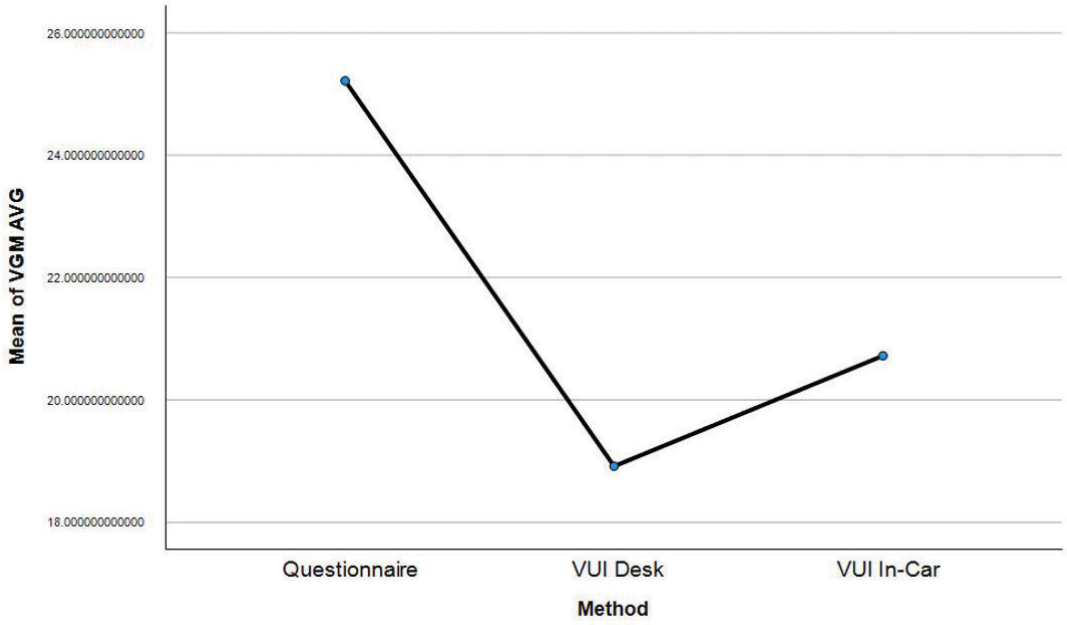


Figure 25 Innovation management score

Individual contrast comparisons explored the significant effects in detail. For Fluency, more ideas were generated in the VUI desk method ($t(55) = 2.1x$, $p < .02$ one-tailed) and VUI In-Car method ($t(55) = 1.8x$, $p < .04$ one-tailed) compared to the questionnaire method. A similar pattern is observed for Flexibility, where the VUI desk method ($t(55) = 1.91x$, $p < .04$ one-tailed) and VUI In-Car method ($t(55) = 2.3x$, $p < .04$ one-tailed) resulted in more diverse ideas compared to the questionnaire method.

Second creativity question

For the second creativity question one-way anova showed a directionally significant effect of method on Fluency criteria ($F(2,66)2.55$, $p < .09$), a significant effect of method on Flexibility criteria ($F(2,66) = 3.56$, $p < .04$), no effect of method on Originality criteria ($p > .11$), and a significant difference between methods on the Innovation management creativity score ($F(2,66) = 3.27$, $p < .05$).

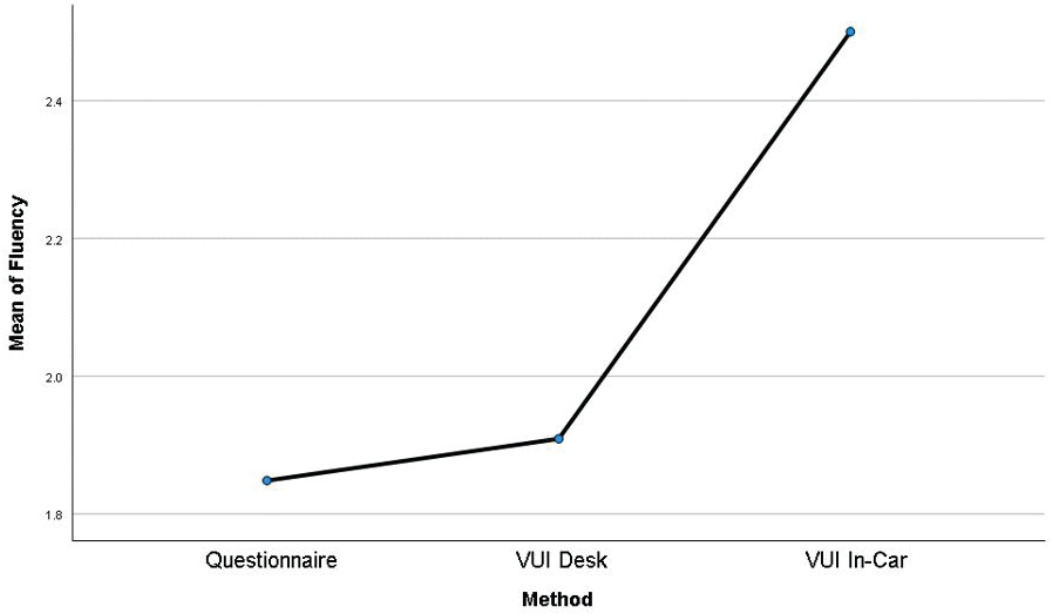


Figure 26 Fluency

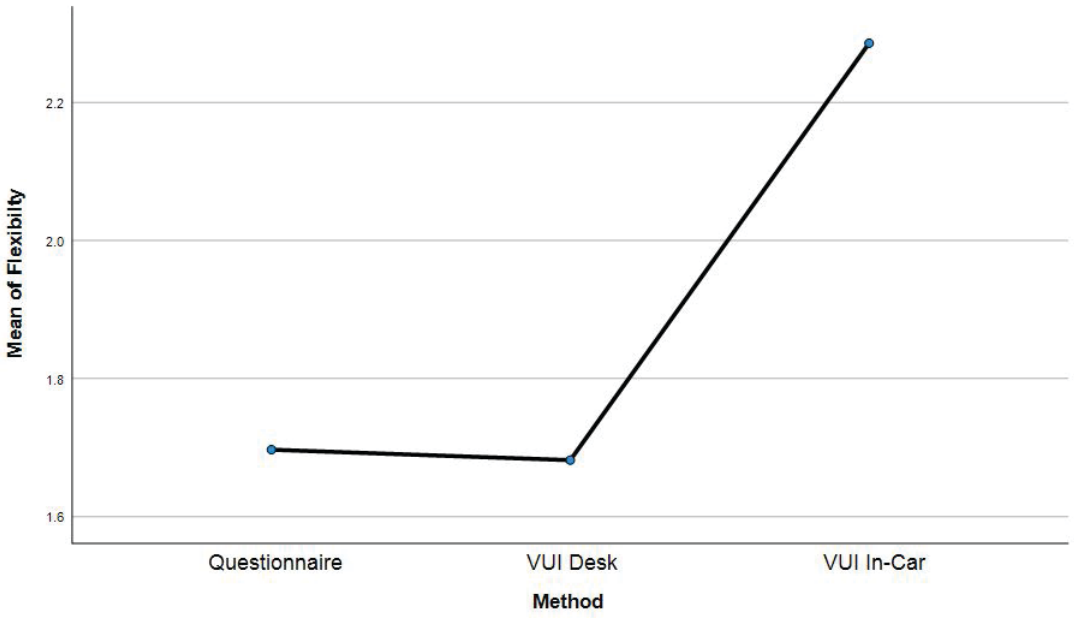


Figure 27 Flexibility

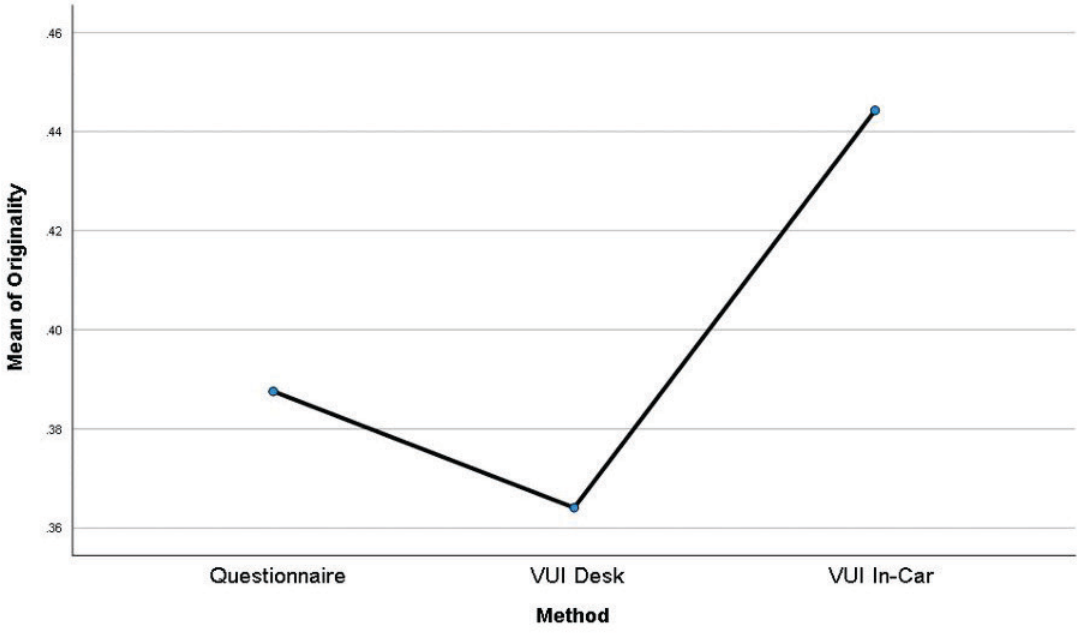


Figure 28 Originality

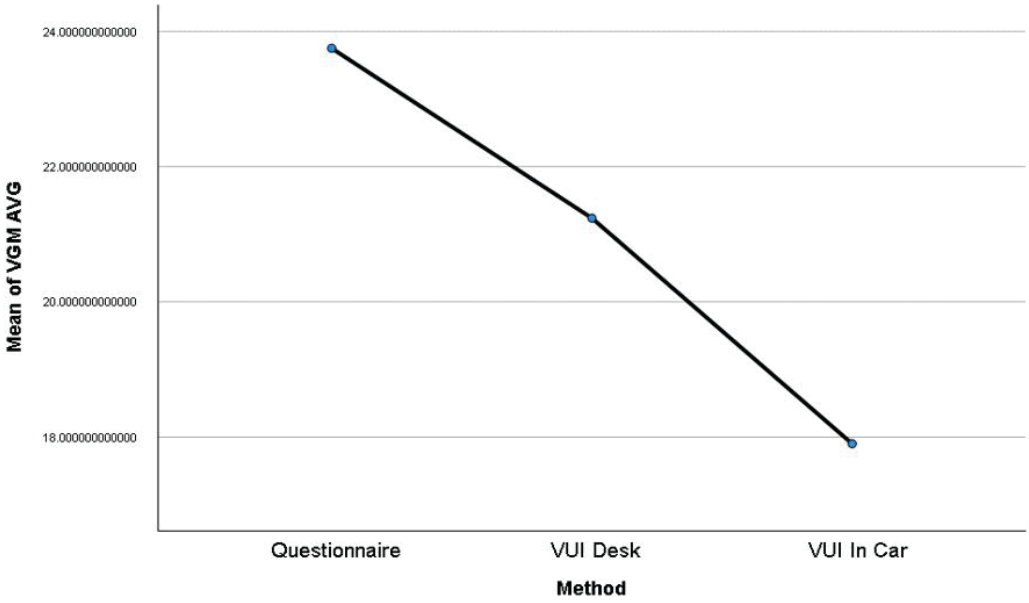


Figure 29 Innovation management score

For the second creativity question, individual contrast comparisons explored again the significant effects in detail. For Fluency, more ideas were generated in the VUI In-Car method ($t(68) = 1.9x$, $p < .04$ one-tailed) compared to the questionnaire and VUI desk method. A similar pattern is observed for Flexibility, where again the VUI In-Car method ($t(68) = 2.4x$, $p < .02$ one-tailed) resulted in more diverse ideas compared to the questionnaire and VUI desk method. Unfortunately, for the Innovation management creativity score, a lower score was detected in the VUI Desk method ($t(68) = 2.5x$, $p < .01$ one-tailed) compared to the Questionnaire. Visually in the mean plot, this also seemed for the VUI In-Car method, but this effect was not significant.
Interestingly, the positive effect of the VUI desk method on Fluency and Flexibility is not observed after this task. In other words, there is a significant difference detected between the VUI desk method and the VUI In-Car method. For all these relations this means that there is an increase in the number of ideas and diversity, which can be visually seen in the plots and appendixes 1.8.1, 1.8.2.

Research Question 2

The results that enable the argumentation on RQ2 are derived from the question where the participants answer to what extent they were able to influence BMW product development. How this was asked of the participant, has been explained in the `VUI customer test´ chapter.

To answer if the participants experienced a higher ability of co-creation after one of the three methods (RQ2), the co-create measure was examined with a one-way ANOVA test. The effect of the method on perceived co-creative contribution to BMW developments was marginally, yet directionally significant ($F(2,78)=2.58$, $p > .08$). However, for this research, a 3% acceptance of not having an actual difference is considered justifiable. Post hoc analysis of individual contrasts showed that perceptions of co-creation are higher for the VUI desk method and the VUI In-Car method compared to the questionnaire ($t(78) = 1.7x$, $p < .05$ one-tailed) and ($t(78) = 2.0x$, $p < .03$ one-tailed) respectively

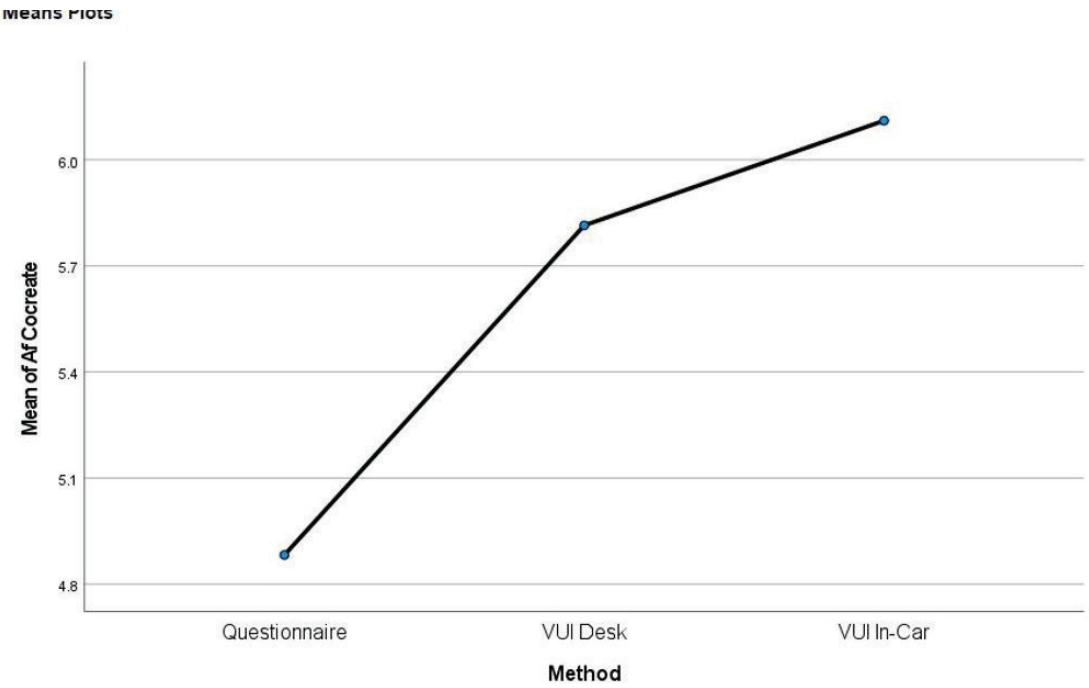


Figure 30 Co-create

Research Question 3

Argumentation to answer RQ3 results from creating a time and cost balance overview between current user tests and the VUI In-Car method, the needed insights on current questionnaire proceedings are gained from the expert interviews.

	Current short term user tests	VUI In-Car method
Preparation time	3-6 months	1-2 months
Analysing time	3-4 months	1-2 months
Amount of customers	n= max 100 per test	n= All modern BMW owners
Data	Qualitative + Quantitative	Qualitative + Quantitative
Data storage	Automated + Manual reflection	Automated speech2text
Time (planning)	Fixed, appointment	Flexible, fitting to users
Employees	5 per team	1-2 per team
Costs	100%	33% total

Table 5 Time and cost summary

RQ3 is being answered by the fact that the VUI IN-Car method improves the data logging and traceability of the data. The current source of occurring problems experienced by general data analytics lies in the fact that feedback is given at later moments in time than when the issue actually happened. The source of the user interaction issues can now be directly linked to the spoken comments of the test persons.

Conclusions

For the Fluency and Flexibility criteria, the differences in results between the first creativity question and the second can be explained by a fixation on contemporary speech problems and solutions in the first creativity question. For the second creativity question, context stimulation in the VUI In-Car method does seem to affect the results. Therefore for the first creativity question results show an increase in ideas and diversity of ideas when the speech product is being used as a tool to perform a feedback session. The product seems to enable a higher creativity result, from the Fluency and Flexibility criteria. For the second creativity question, the In-Car context seems to enable the participants to generate a higher amount and more diverse ideas.

The fact that in the first creativity question the creativity rating from the Innovation management analysis value does not increase from the questionnaire to the VUI Desk of the VUI In-Car method, can be explained by the possible fixation on contemporary speech problems and solutions. The question is aimed at current problems and therefore participants are fixating on existing problems and coming up with solutions that possibly solve them, it appears that using a VUI as a tool combined with questions about this tool, does not result in more creative input.
For the second creativity question, this Innovation management analysis can be seen that there was a decrease in creativity from the questionnaire to the VUI In-Car test. This could be explained by the fact that the prototype had some minor technical setbacks and therefore could distort this test. This will be assessed in the discussion.

For the Co-create values, the data showed that from the questionnaire to the VUI Desk test and from the VUI Desk test to the VUI In-Car test there is a positive relation. As shown in the mean plot in appendix 1.8.4 the biggest step is made from the questionnaire to the VUI Desk test. The fact that participants actually interact with a speech product while giving feedback on the topic of speech interaction in a car, does result in a higher co-creation factor.

My tested VUI IN-Car method has great time saving potential compared to contemporary customer feedback methods. This benefit can be explained because of the current missing directly recorded link between; locating actual problems and pinpointing the problems users address in contemporary feedback methods. Due to the fact that these contemporary customer feedback methods enable product feedback that is being given at a later moment than the actual problem occurred. Therefore, the analysing time for linking customer feedback issues to the reason for the issue is part of the current data analysing. This is step is reduced almost entirely when a VUI IN-Car method would be implemented, because of individual real-time connections between the issue and the reason for the issue.

Discussion

Some technical limitations from the created methods and also the execution of the tests were taken into consideration in this summary of the results of this research.

Technical limitations from the VUI. However, the time paced factor has been reduced as far as possible. Still, the fact that the questionnaire was self time-paced and both the VUI tests remained machine paced. The time available for answering the questions was tested during the pilot expert testing phase, still, this was not individualized to each participant during the actual data gathering. This will have had an effect on a part of the participants. Some participants could have used more time and others on some occasions needed to wait for the VUI to proceed.

For the VUI tests, I needed to be present in the room where the tests took place. The awareness of someone present that is running the tests can have an influence on the participants. This is something further research could dive into and propose future solutions if needed.

To an extent, there has been a bias factor in my tests, mainly because the participants will have affection for the brand `BMW´, "since the participants are being employed by BMW and therefore represent the brand" (Expert, 15). At the current pre-development stage of this research, this is a limiting factor that has to be accepted for confidentiality reasons. However, according to literature testing with experienced users, can be a solution to prevent design fixation on examples in customer studies (Schoormans, Ortt, and De Bont, 1995; Von Hippel, 1986). And therefore is not necessarily result in a negative influence on the data in this research.

Pace related results show that some participants in the questionnaire took up to a factor of 3 more time than the participants in both of the VUI tests had. On average participants in the questionnaire session took twice as long to answer the same questions. This will have had an effect on the creativity values. Therefore, time limitations could have been addressed for the questionnaires to generate better matching circumstances. The VUI methods have been addressed in order to operate similar to a questionnaire, but from the questionnaire perspective this should need an extra step in terms of time pacing.

Considering these factors, the results from this research are very promising. Even from this developed prototype statistical evidence points out that VUI methods can be beneficial for creative customer feedback tests. Therefore, I have looked into potential implementation possibilities for BMW.

BMW GROUP

"I am convinced such a tool can solve problems we have had for 20 years".

"Sure the internal testing departments will be interested in your tool, commenting live on errors or user issues while driving is not only faster and better traceable but also safer than the current procedures".

8 - Roll-out Strategy

8 - Roll-out Strategy

Roll-out Strategy

From these results, a roll-out strategy will evolve, that will enable developers to have early feedback on their ideation or visions in the pre-development stages. Low-fidelity concept projects can be validated with quality feedback from customers and will touch upon the potential pitfalls to avoid in involving customers in pre-development stages. Therefore, work towards an agile way of gathering consumer feedback, before preparing the larger user testing phase starts.

The Head of the UI/UX department (Expert, 27) sees the potential of gathering direct and traceable customer feedback via speech, "I am convinced such a tool can solve problems we have had for 20 years". This interview with the highest authority opened doors to more experts. User testing department (Expert, 24) inspired the implementation possibilities that formed the first rollout implementation in the BMW organization: "Sure the internal testing departments will be interested in your tool, commenting live on errors or user issues while driving is not only faster and better traceable but also safer than the current procedures".

In the development phase after pre-development, the series-development phase follows. This is a phase where certain expert employees of BMW are able to test cars over the de weekend and report issues and findings afterwards in an app.

The implementation can be made with a small effort by implementing the VUI In-Car method in the discussed series-development phase. Where the discusses expert user tests take place. Currently, my Alexa skill is already able to run on series production cars. Therefore, the first implementation tests are already implementable in the discussed series-development employee tests. The benefit for the targetted departments is mainly the time-saving aspects on the data linking between issue and documentation of customer feedback on the issues.

The next step would be testing and selecting the most suitable software back-end to run this feedback session on. First iterations are possible via the same approach as I performed the tests. However, the aim would be to run the feedback sessions on more permanent software in order to enable scaling up this feedback method to multiple departments and purposes.

Concluding, from the interviews performed with the corresponding departments implementation of the tool that resulted from this research is enabling faster, safer and better traceable feedback in internal user testing. By installing speech interaction feedback as a tool to report user feedback.

9 - Sources

Amabile, T. (1996). Creativity in context. Boulder, CO: Westview Press.

Carbon, C. C. (2015). Predicting Preferences for Innovative Design: The "Repeated Evaluation Technique"(RET). NIM Marketing Intelligence Review, 7(2), 34-39.

Carbon, C. C. (2019). Psychology of design. Design Science, 5.Cheng, P., Mugge, R., & Schoormans, J. P. (2014). A new strategy to reduce design fixation: Presenting partial photographs to designers. Design Studies, 35(4), 374–391.

Carbon, C. C., & Leder, H. (2005). The repeated evaluation technique (RET). A method to capture dynamic effects of innovativeness and attractiveness. Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 19(5), 587-601.

Cheng, P., Mugge, R., & Schoormans, J. P. (2014). A new strategy to reduce design fixation: Presenting partial photographs to designers. Design Studies, 35(4), 374-391.

Gonçalves, M., Cardoso, C., & Badke-Schaub, P. (2013). Inspiration peak: exploring the semantic distance between design problem and textual inspirational stimuli. International Journal of Design Creativity and Innovation, 1(4), 215–232.

Guilford, J. (1950). Creativity. American Psychologist, 5, 444–454

Hair, J. F., Barth, K., Neubert, D., & Sarstedt, M. (2016). Examining the Role of Psychological Ownership and Feedback in Customer Empowerment Strategies. Journal of Creating Value,2(2), 194-210. doi:10.1177/2394964316661811

Hogenhuis, B. N., van den Hende, E. A., & Hultink, E. J. (2016). When Should Large Firms Collaborate with Young Ventures? Understanding young firms' strengths can help firms make the right decisions around asymmetric collaborations. Research-Technology Management, 59(1), 39-47.

Hunt, D. M., Geiger-Oneto, S., & Varca, P. E. (2012). Satisfaction in the context of customer co-production: a behavioural involvement perspective. Journal of Consumer Behaviour, 11(5), 347-356.

Opata, C. N., Xiao, W., Nusenu, A. A., Tetteh, S., & John Narh, T. W. (2020). Customer Value Co-Creation in the Automobile Industry: Antecedents, Satisfaction, and Moderation. SAGE Open, 10(3), 2158244020948527.

Poetz, M. K., & Schreier, M. (2012). The Value of Crowdsourcing: Can Users Really Compete with Professionals in Generating New Product Ideas? Journal of Product Innovation Management, 29(2), 245–256.

Razavi, M., Adavi, S., Alam, M. Z., Mohr, D., & Zachmann, G. (2014). Innovative and Contact-free Natural User Interaction with Cars. In Proceedings of the International Conference on Virtual and Augmented Reality (GI VR/AR Workshop).

Schweitzer, F., & Van den Hende, E. A. (2017). Drivers and consequences of narrative transportation: Understanding the role of stories and domain-specific skills in improving radically new products. Journal of Product Innovation Management, 34(1), 101-118.

Schweitzer, F., Rau, C., Gassmann, O., & van den Hende, E. (2015). Technologically reflective individuals as enablers of social innovation. Journal of Product Innovation Management, 32(6), 847-860.

Schweitzer, F., Van den Hende, E. A., & Hultink, E. J. (2019). There's more than one perspective to take into account for successful customer integration into radical new product innovation: A framework and research agenda. IEEE Transactions on Engineering Management, 67(3), 813-829.

Sio, U. N., Kotovsky, K., & Cagan, J. (2015). Fixation or inspiration? A meta-analytic review of the role of examples on design processes. Design Studies, 39, 70-99.

van Boeijen, A., Daalhuizen, J., Zijlstra, J., & van der Schoor, R. (2013). Delft Design Guide: Design strategies and methods. BIS Publishers.

van Den Hende, E. A., & Schoormans, J. P. (2012). The story is as good as the real thing: Early customer input on product applications of radically new technologies. Journal of Product Innovation Management, 29(4), 655-666.

Ward, T. B. (1994). Structured imagination: The role of category structure in exemplar generation. Cognitive psychology, 27(1), 1-40.

West, J., & Bogers, M. (2014). Leveraging external sources of innovation: a review of research on open innovation. Journal of product innovation management, 31(4), 814-831.

Yao, X. & Vink, P. (2019). Design for a Better Hygiene Experience in a Flying-V Aircraft.

BMW Natural User Interaction (NUI)* at the Mobile World Congress 2019 in Barcelona