

DCB412: DESIGNING FOR THE USER EXPERIENCE

Arthur Geel, Selim Haase, Rosa van Koningsbruggen

TABLE OF CONTENTS

1.	Economic paradigms and human needs	02
1.1	Introduction	02
1.2	Three applications of music	03
2.	Design-relevant knowledge in early design phases	07
2.1	Three research methods	07
2.2	Pilot study	11
2.3	Final study	12
2.4	StoryPly	14
3.	User experience evaluation	16
3.1	Various questionnaires	16
3.2	Individual homework assignment analysis	18
3.3	Statistical analysis of questionnaire	20
4.	Personal reflections	23
4.1	Selim Haase	23
4.2	Arthur Geel	24
4.3	Rosa van Koningsbruggen	25
5.	Appendices	26
5A	Task division	26
5B	Questionnaire data	27
5C	User tests and data	29
5D	Summaries	30

Arthur Geel	B2.2	s149144
Selim Haase	B2.2	s147431
Rosa van Koningsbruggen	B2.2	s132640

1. Economic paradigms and human needs

1.1: Introduction

This elective started off by introducing us to what user experiences actually are, and how the basic human needs are related to this. As stated in the first lecture, an experience may be considered as “*a story that emerges from the dialogue of a person with his or her world through action*”¹. Building on that, we learnt about the economic paradigms, where the basic human needs are regarded in the four distinct paradigms. These human needs are *Autonomy, Competence, Relatedness, Stimulation, Popularity and Security*.

Our general theme for the applications of products in an economic paradigm context was music. In the upcoming chapter we will be discussing three applications: the *Sony Walkman WM-2 (1981)*, the combination of *iPod and iTunes (2005)* and the *Spotify* music streaming service.

These three applications were selected because they arguably caused some of the most drastic developments in the context of music, which will be touched on later. The applications are listed in chronologic order. As a start, these three applications have been mapped to the basic human needs, see fig. 1.

Fig. 1: Three applications of music, mapped to the basic human needs.

	Sony Walkman	iPod/iTunes	Spotify
Autonomy	Freedom, independence	Freedom, independence	Freedom, independence
Competency		Control, performance	Control, performance
Relatedness			Presence, emotional expression
Stimulation	Novelty	Novelty, play	Curiosity
Popularity	Status, recognition, fashion	Status, recognition, fashion	Recognition, helping
Security	Relaxation	Relaxation	Relaxation

¹ Hassenzahl, M. (2011). DCB412: Designing for the User Experience [Powerpoint lecture]. Retrieved from <https://dlwpswbsp.tue.nl/190-2015/744a2ce9e4ae4022968358afb26e7827/Documents/DCB412%20Lecture%201.pdf>

1.2: Three applications of music

Sony Walkman WM-2: Experience paradigm

The Sony Walkman MK-2 was one of the first portable music players. It allowed people to take cassette disks of their favourite music with them, and listen to it (privately) whenever they wanted to. It is unique in the sense that it broke quite some social taboos. At the time of the walkman it was considered rude to wear headphones in public, as it was related to being excluded of the social environment. The walkman has helped the acceptance of headphones in public spaces.

It also was a huge breakthrough brand marketing. As Sony was a fairly established company already, they had the power to create their own brand experience. The production of the walkman by Sony was a major key in its success. In order to keep their customers coming back, they created music labels, which then again only created music that can be heard from a Sony device. This made the product shift from product experience to brand experience.

Experience economy	
Experience	Marketing & branding
Global	Brand experience
Explore lifestyle identities	Product-service mix
Work hard play hard	Targeted experiences
Experimentation	Promote brand lifestyle
Break social taboos	Growth

Finally, Sony offered a product-service mix. The walkman was one of the first products to standard come with a global warranty. This meant that the users were not just buying a physical product, they were buying a service to come along with it. It is because of these reasons that we have concluded that the Sony Walkman MK-2 belongs to the Experience Economy (Class II).

iPod/iTunes: Experience & Knowledge Paradigm

When Apple launched their iPod nano in 2005 they changed the way music worked. The iPod nano was a music player capable of holding up to one thousand songs at the time of release. This was something that had not been done before. Because of this, people were no longer limited to the amount of music tapes they brought with them; all of their favourite songs were now available anywhere they wanted.

The most innovative trait of the iPod nano was the seamless interconnectivity with Apple's music service, iTunes. The iPod nano required this software to be used in order to upload songs to the device, so the service had a substantial group of users. The main feature of iTunes was the possibility for users to easily buy music online. Instead of having to buy entire albums just to access a single song, iTunes allowed buying single songs for a modest price. In general, Apple are very capable with their brand lifestyle approach, and the iPod nano is no exclusion. The fact that iTunes allows users to listen to music they have bought on different platforms (Music players, personal computers, etc.) strongly hints that this product is

part of the Knowledge Economy (Class III), yet other parts still are in the Experience Economy (Class II).

Spotify: Knowledge Paradigm

Spotify is one of the leading legal online music streaming services. Since its launch in 2008, it provides two 'user experiences': *freemium* and *premium*. It is one of the first platforms to offer free, legal music to its customers, which made music a whole lot more accessible to everyone. This is one of Spotify's key assets: by using this service you can easily access all kinds of music without having to pay (a lot) for it. Instead of the old model where you had to buy entire albums just to be able to listen to a single, with Spotify you can find countless of songs in the library to enjoy within minutes.

It's not just easy to find the songs in the libraries, the songs can be taken anywhere. Spotify has a mobile application which allows you to bring your music with you while you are on the go, which gives the user a lot of individual empowerment. As a music consumer, you no longer need to think about what music you need to bring before you head out.

Finally, Spotify has a lot of value if you consider the amount of data that is being produced when people use it. In order to use Spotify, you need to create a individual account, which has access to some of your personal data. Because Spotify is mainly used in combination with an active internet connection, this means that this data can be analyzed to offer the users more value. Artificial intelligence allows the system to easily suggest new songs that the users will most likely enjoy. The analysis of the data can also be used to connect users of Spotify, allowing them to appreciate music in a social setting. Unfortunately, this functionality has not yet been made fully available to the broader audiences, as it currently is only available for software developers. However, we expect this to change within the next few years, as this can be a huge add-on to the current product.

Conclusion:

Looking back at the three highlighted examples of music applications we can deduce that the more advanced products become, the more human needs it needs to satisfy in order to become an economic success. In the Industrial Economy, products were simply manufactured and sold to users, as the main focus of the products was on the functionality.

In the Experience Economy, companies were eager to include a service with their product. Instead of stopping contact with users after the sale, companies offered perks and benefits to users for using their products. This led to products favouring a specific brand, also known as the brand lifestyle. This is something beneficial to both consumers and companies: consumers get a better, standardized experience while companies have a steadier group of customers, and a generally more stable economy.

Finally, in the Knowledge Economy we can see a shift where the users change their stance from passive to active. Something companies have started implementing in this economy is the base for users to indulge in their creativity. Users are stimulated to participate to add more value to the services themselves.

In the following subchapter the three aforementioned applications of music will be observed on potential points that can be improved, to 'upscale' these applications to a next economic paradigm.

Sony Walkman WM-2 (Selim)

How can we take the Sony Walkman to the next level?

At the time the walkman was released there was no such thing as digital playlists. Therefore such a function will increase the personalisation of the product. Therefore it will empower the user to create his/her own playlists by making it much more convenient.

Also to shift it to the next paradigm it could be possible to create a sense of community instead of branding. This community could be created around a sharing platform that is an extension of the walkmans. People might like others playlist and share them too. This might enable a shift out of the experience paradigm.

To enable a shift to the knowledge economy, the walkman could have functions that enable the user to create value for the product itself. This can be achieved by creating an open source music sharing platform where the users themselves can upload songs that they have create. As result, the brand itself will be more in the background and the users will be in front since they are able to create value in combination with the brand.

iPod/iTunes (Rosa)

How can we take the iPod/iTunes combination to the next level?

In order to upscale the iPod/iTunes combination Apple needs to become more open. Their product themselves are already leaning towards the Knowledge Economy, as they allow individual empowerment, creativity, and the view can be described contextual. However, the business mind-set of Apple is fully focused on the Experience Economy. This could be changed by creating open tools and allow the customers to participate and self-develop. Apple could allow people to customize the operating system, or co-design aspects of their products and services.

Nevertheless, it is a very deliberate decisions of Apple not to do such a thing. By keeping their business mind-set in the Experience Economy, they create a closed and controllable environment, and distinguish themselves from Android. This allows them to create a very strong brand lifestyle, as restricting certain content obliges users to use other products of the Apple brand to ensure a smooth interconnectivity between devices.



Spotify (Arthur)

How can we take Spotify to the next level?

In order to upscale Spotify we need to focus on making the service more social. A major part of the transformation paradigm is addressing the social- and environmental problems, which is something that is severely lacking at the moment. With Spotify, a musical network was created that has the potential to be so much more. Application developers for Spotify already have access to a lot of things that allows them to gain more insight in how people use music in a social way, yet hardly any of these features are available for the general public. An example of functionalities like these are the availability of detailed statistics of the audience, like gender, age, location and so on. This yields the potential to make Spotify into a social network where like-minded individuals can organise and attend specific events.

A step in the right direction would be to give music fanatics the opportunity to create a strong community where they are encouraged to share their preference in songs, artists and albums. A big part in the transformation economy is a feeling of empathy and cooperation, which is something that Spotify lacks at the moment. If people were given more opportunities to create a community, this would be achieved.

Still, it is doubtful if these additions to Spotify would make it eligible to be considered an example of a Transformation Economy. As we have experienced in the lectures, it can be hard to find good examples of a Transformation Economy, simply because it still is a novel concept, and it is expected to make a bigger impact in the future.

2. Design-relevant knowledge in early design phases

2.1: Three research methods

For the following weeks we were asked to look into three methods to gain design-relevant knowledge in early design phases: *Cultural Probes*, *Diary Studies* and *Contextual Inquiries*. The following are our individual insights on these aforementioned methods. At the end of this subchapter we will compare the methods and draw conclusions based on that.

Cultural probes method (Arthur): This is a unique way to gain inspiration and to generate ideas with the help of users. It usually consists of a package of tasks for the participants to do in their daily lives, which documents data that can be analysed by the researchers later on.

The main advantage of this method is that it's a novel and fun way to conduct a research. Often, research is seen as factual and monotonous, so it can be hard to find new insights because of uninterested participants. Because the cultural probes consist of thought-provoking items, participants usually are more interested in giving detailed answers.

The main disadvantage is that the data gathered is very subjective, so it is hard to get concrete results that can be used for deeper analysis. Therefore, a method like this should primarily be used in studies that are generative, rather than evaluative.

We used this method in this elective. We created an interactive questionnaire with a number of novel assignments that the participants had to complete. These assignments were implemented to help keep the participants eager to give more in-depth answers to help us assess their context, problems and opportunities.

Diary studies method (Arthur): This is a way to gather data by analysing self-reported data of participants. The participants are asked to document their lives for a set amount of time, which will later be analysed by the researchers.

The main advantage of this method is that the data gathered can provide a clear image of what the motivations for participants were. It can provide a very clear timeline where the researchers can see what happened before, during and after the main events. Because of this, problems and opportunities are easy to locate.

The main disadvantage again is that the generated data is subjective. Participants can be careless with their documentation of the events, leading to inaccuracies. Also, because participants are asked to document it themselves they can give their own interpretation of the events, which can lead to incorrect biased tendencies.

A prime situation where this method can be used is for a generative research to locate problems in users' daily lives. We could have used this method in the context of travelling with music. We wanted to find ways to improve the user experience, either by making operating the music more pleasant or by taking away stress. The participants of a research

like this could have been equipped with a device where they can indicate stress/anxiety, coupled with a diary that is to be documented after they finished travelling.

Contextual inquiry method (Arthur): This is a tool to help you determine in detail what the *actual needs* of users in a particular environment are. This method is a comprehensive approach that requires the researchers to actively study the participants through inquiries, observations and interpretations.

The main advantage of this method is that it yields very reliable data. Because of the extensive work that the researchers put in, the data is eminently detailed and accurate. Also, this method has proven to be very flexible, so it can be used in most contexts. Therefore, research done through contextual inquiries often provide valuable new insights in those conditions.

The main disadvantage of this method is that it's very demanding, for the participants but mainly for the researchers, as they have to be near the participants whenever new data is to be gathered. Out of these three methods, it's the most time-intensive method. Because of this, it is hard to get high sample sizes of users, which makes this method prone to statistical inaccuracies.

Cultural Probes (Rosa)

Cultural Probe is a very creative method that allows the researcher to get very personal insights. It is also a way to explore many different aspects. Another advantage of Cultural Probes is that the data is reliable, as the users are the experts of their own lives and the data is independent (no one interferes with the user, which can be the case with other methods).

However, there are also drawbacks to this method. The amount and quality you receive is not certain. Some people might not take it seriously or misinterpreted what they have to do. Furthermore, it is time consuming and geographic/cultural distances might be a problem. In order to have a good Cultural Probe, one had to minimize the distance between the researcher and the user by creating a personal relationship.

For me, it would be useful to use the Cultural Probe method in order to get an idea of what the problems are. You receive very personal data, which gives you a clear image of user needs.

Contextual Inquiry (Rosa)

Contextual Inquiry is a good method to reveal knowledge about the user's work, that the user himself/herself is not aware of. It also gives the researcher an idea of the environment and what the user truly needs. Lastly, this method provides detailed information and it is a flexible technique.

However, it is also an intense method and the researcher needs to be familiar with the domain, in order to obtain valuable data.

I could see myself using this method when want to understand the environment of the user, so that I know what they need. For example, last year I was working on a project called the Plan Gem. For this project, we made an agenda/calendar that you could feel. The idea was that eventually, the user would start to 'feel' time, creating a 6th sense. For the Plan Gem, it would have been a good user test to have a Contextual Inquiry that looked into what information people want to know about their agenda in order to work efficiently.

Diary Studies (Rosa)

One of the main advantages of a Diary Study is that it does not cost the researcher much time, as the users are capturing the events, rather than the researcher. The users can use photos and audio to recall events more vividly and this gives the researches a better idea of the everyday interactions. Lastly, the users are not constantly interrupted.

However, the participants need to be disciplined and there is no way to verify the results. Also, the researcher does not have the chance to ask questions during the interactions, only afterwards.

I think this method would be useful when you have a minimal viable product. You could give this product to the intended user group and receive data on how they use it, what they think of it, and what they would improve.

Cultural Probes (Selim)

I believe cultural probes to be a fun and empowering method. It can create an enjoyable atmosphere through its somewhat informal approach which can also make the user more comfortable. Also the data is very personal and unbiased since there is no direct interference between users and the researcher.

On the other hand, the amount and quality of the data is uncertain. It also requires a lot of focus and effort of the user. Because of this the user might lose his attention and results might become sloppy.

When you want to understand a context, cultural probes can give you personal insights that are unbiased. For example, if I would like to understand the context of working from home, and I don't know much about it, I could use cultural probes to gain insights from people that live daily in that context.

Contextual Inquiry (Selim)

Contextual inquiries can make you understand the environment of the user and understand their (latent) needs. It can make you aware of aspects that user isn't aware of himself.

In order to use the method the research should be familiar with the research domain in order to use effectively.

To understand the environment of the user, so that you know what they need. For example, you want to know what information people want to know about their agenda in order to work efficiently.

Diary studies (Selim)

This method is very time efficient. In order to remember experiences better recording media, such video and photos can be used. This can also be done by the users themselves.

One drawback is that there is no way to verify results gained by this method.

It can be used to receive data on how the users feel and think about a design.

I would use this method to identify problems that are in the daily lives of users. Even if they aren't aware of them, the researcher might become aware through diary studies.

Conclusion

The three aforementioned research methods are quite diverse, and can be used for various purposes in different stages of design processes. A method hardly ever is 'wrong' in a certain context, sometimes it just is slightly more beneficial to use one research method over the others because of small differences.

The main thing we have learned from examining these three research methods is their area of use. The three methods have huge differences in how much effort they require, both for the ones doing the research as the ones merely participating in the research. The Contextual Inquiries require the most work by a great deal. Since it also requires the researchers to be present during interviews with the participants, it is very resource-heavy to do multiple Contextual Inquiries at the same time. The results you do get from Contextual Inquiries are often regarded as scientifically valid though, which is why this method is mainly used in later stages of design processes.

Diary studies however require far less input from the researchers. The drawback of this method however is the fact that the effort has to come from the participants, which makes it possible for the participants to produce results that have little to no value. The results also carry less academic weight, other methods can be more credible and convincing.

While cultural probes require some more effort than diary studies, they often are seen as more engaging for the participants, leading to more valuable results. The results however are generally quite shallow, which occasionally makes this method non-favourable for deeper analysis. Instead, a different research method should be used.

2.2: Pilot study

Method used

For the research we chose to work with the cultural probe method. This method would allow us to get a general personal insight in people's music listening experience on the go and as music is a personal experience, we believe that this method was best suited because it allowed us to expand the context to get more information in the time-sensitive period we were facing. Also, the idea of using cultural probes seemed valuable to us in the sense that it would stimulate the participants to produce worthwhile conclusions.

Case study

Almost everyone in the Netherlands uses a bike regularly and a lot of people want to make the bike ride a more pleasant experience by listening to music whilst cycling. However, do we feel and are we indeed safe listening to music and cycling at the same time?

This will be examined with the cultural probe.

Design goal

The cultural probe will be conducted for the following design goal:

"We want to make listening to music whilst cycling a safer experience."

Research question

"How can we understand whether people who are cycling want to listen to music and whether they feel comfortable?"

Pilot Study

For the pilot study all the team members had to conduct a probe with a small sample size to check if the results of the actual research will be useful. This study was held over four participants: two female, two male, aged 18 to 24. Each participant had to fill in some general information about themselves which was done in order to get an overview of the demographics. After having done that, the participants had to answer some closed questions about the time they spend on a bike, whether they listen to music whilst cycling, and after what time they start to listen to music. This would give us an indication whether people even listen to music when cycling and for how long they have to cycle before starting to listen to music.

The next part of the probe had to do with negative experiences the participants possibly had when cycling and listening to music. They each had to make a sketch of the situation and describe this situation, in order for us to get an image of when people do not feel safe. The ultimate goal of this would be that we can map the situation that are regarded unpleasant according to the participants and design something that would improve situations.

Furthermore, all participants had to draw their route on a map. The goal of this assignment was to get an idea of the duration and 'dangerous' points the participants encountered on their daily route.

Lastly, the participants had to colour in some bars that would give an indication how people listen to music whilst cycling. So what volume they listen to, relaxed or energized music, and whether they used playlists, shuffle, or albums.

From the pilot we could conclude that people start to listen to music when their trip is ten minutes or longer and all participants have experienced a dangerous situation whilst cycling. This situation was in all four cases a scooter driving behind the cyclist and causing a dangerous situation.

The last part of the probe led to the conclusion that people use shuffle and playlists, listen to the music on a volume of approximately 75%, and the music tends to be more energizing.

However, as this was a pilot study, there was still a lot of room for improvement. From the pilot we learned that we wanted to include more sketches and photographs in the real study, that we had to add more specific questions on dangerous situations, and with the bar section add bars for listening to music whilst walking, in order to see whether there is a difference between music listening whilst cycling and whilst walking.

2.3: Final Study

After the pilot study we made some improvements to the cultural probe. More detailed questions on safety and listening to music whilst cycling were asked. Furthermore, we included more sketching areas, due to the added questions.

We elaborated on the closed question part of the probe, by asking the participants if they listen to music whilst cycling and whether comfortable when cycling, without and with music.

We also added an extra bar section for listening to music whilst walking. This way we could compare whether there would be a difference between the music experience for walking and cycling.

Lastly, we added a legend to the map of Eindhoven, so that people could draw multiple routes on it. The cultural probe can be found in Appendix C on page 28.

For this actual probe we had a total of nine participants. The demographics were the following:

- 6 male
- 3 female
- Aged 19- 22
- 8 students
- 1 unemployed

From the closed question part it was noticeable that most of the participants do feel safe when cycling with music. Before this study, we assumed that people would be slightly more uncomfortable when listening to music. However, with this study, it appeared that most of them are fully comfortable.

It was also interesting that all the participants listen to music when cycling from time to time, which gives us a nice target group. Also, the participants start to listen to music when the trip is ten minutes or longer.

The open questions and sketches provided a good insight in what the participants experienced as unpleasant and dangerous. Rather often it was stated that they themselves created the dangerous situation by looking at their phone and not paying attention to the surroundings. Another situation that was often given, was that the participants could not hear the traffic as good as without music. Lastly, the participants named cycling through dark forest at night cycling along a highway, and cycling through crowded traffic situations as unpleasant.

Analysing the routes did not provide us with the information we hoped to receive. However, this is valuable knowledge for future cultural probes or future iterations on this probe.

The last part that had to be analysed was the part of the bars. Whilst analysing this, we could conclude that there is not much of a difference between listening to music when walking or when cycling. Most of the cases, the participants had the exact same listening preferences. In a few cases, the participants listened to more relaxing music when walking.

From the final cultural probe the following conclusions could be drawn: generally speaking, people have the feeling that they are safe listening to music whilst cycling. People start to listen to music when their trip is ten minutes or more. And there are no significant differences between walking and cycling. Some people did listen to more relaxing music whilst walking and more energetic music whilst listening. The routes were not really of use to us and we could not draw a conclusion from that part.

The following conclusions have given us a better idea of how people listen to music on a bike and whether they feel comfortable. The probe concluded that people do feel safe listening to music whilst cycling, which answers one part of our research question: *“How can we understand whether people who are cycling want to listen to music and whether they feel comfortable?”*

The other part of the research question can also be answered with the probe. According to the closed questions, all participants like to listen to music when cycling, especially when a trip becomes more than 10 minutes.

If we were going to iterate this probe in order to improve it, we would like to further define what ‘safe’ is and what factors contribute to this feeling. Once we know that, we could create more specific questions and have more specific answers.

Nonetheless, still a lot of accidents happen involving cyclists and other road users. When we combine this with the knowledge gained from the probe, we see a lot of design opportunities to make it listening to music a safer whilst cycling a safer activity. This will be discussed in the next sub-chapter.

2.4: StoryPly

After having conducted and analysed the cultural probe, there was a workshop in StoryPly. For this workshop we had to select the basic human needs that we were going to address with our concepts. For us they were safety and competency. Safety, as the design goal was to come up with a concept that would allow cyclist to be safer whilst listening to music and competency, as we wanted to enable cyclist to control their music experience, without having to touch their phone.

With the following in mind, we wanted to design something that would make listening to music whilst cycling a safer experience, by giving the user extra competence. In order to do so, we first had to describe the tension between the real situation and the desired situation. So we described the tension as following:

“I want to be safe listening to music and controlling my music whilst cycling, but I am not.”

With this tension in mind, we started to develop the experience story. For the story we had two main characters, the cyclist and one other road user, for purpose of the story we chose this person to be an automobilist.

The cyclist is enjoying her cycling trip and listening to the music. After a while, she becomes bored with the album and wants to switch to something a bit more uplifting. In order to do this, she grabs her mobile phone out of her pocket, unlocks it, and opens the music app. All this time, she is not really focused with the traffic around her and does not notice that she is approaching a dangerous crossroad. On this crossroad she will meet the automobilist, who is not too keen on speed limits. Because our cyclist is still busy with her phone, she enters the cross road and collides with the automobilist.

In order to limit the previous described dangerous situation, we want to create a safer experience. With our concept, the story would go slightly different. Again, the cyclist is enjoying her bike trip and listening to music. However, this time when she becomes bored of the album and wants to switch to something more uplifting, she does not have to grab for her mobile phone, unlock it, and open the music app. Without having to release the steer, she can change the album and volume. This way, she notices that she is approaching a crossroad with the speeding automobilist and brakes in time.

These ‘stories’ provided a clear image of what we wanted to improve and on what point we could improve the experience. Therefore, we came up with 3 concepts:

1. A headphone that adapts its volume to the environment. Once the user approaches a dangerous situation or location, the volumes is turned down, so that the user can focus on the situation.
2. An apparatus that can be tied to the steer of the bike. When the user is cycling and wants to change the album, he or she can make gestures over the apparatus. However, with this solution the user still has to let go of the steer with one arm and gestures are rather distracting, so therefore, we decided that this was not the best idea and improved on it with concept 3.

3. A steer handle on which you can tap to make adjustments. If the user wants to change the volume to max, for example, he or she can give 3 small taps with his or her finger. By allowing user to change their volume and music using their fingers, they won't have to let go of their steer and as these gestures are small and low-effort, they can stay focused on the traffic.

3. User experience evaluation

3.1: Various questionnaires

In our questionnaires we were interested in finding out how people perceive various music streaming applications. Our two applications of choice were Spotify and Soundcloud. The main domains we were interested in were how people perceive the aforementioned applications' potential to help users discover new music and share music with others. Therefore, our research question was:

“How does Spotify compare to Soundcloud in terms of perceived ease of use and perceived usefulness?”

Specifically, we wanted to compare both applications on their perceived usefulness and perceived ease of use.

The questionnaire used for this study is based on the Technology Acceptance Model. This questionnaire is a well-studied and tested instrument. It is also applicable to software which made it a good fit for our research.

For the questionnaire, a total of 18 participants were recruited. These 18 participants were divided in two equal groups so that both applications had 9 participants. This was done to give us a more representative sample size, even though we later concluded that our amount of participants was not sufficient (*see statistical analysis*). The participants were all students and approximately 50% women and 50% men. Their ages were ranging between 19 and 21.

Two other questionnaires (Rosa)

UTAUT

Paper read: V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: Toward a unified view.," MIS Quarterly, vol. 27, pp. 425-478, 2003)

This is basically an improved version of the TAM model. It explains user intentions to use an information system and subsequent usage behaviour. The UTAUT model looks into the following aspects:

1. performance expectancy (the degree to which an individual believes that using the systems will help him/her to attain gains in job performance)
2. effort expectancy (the degree of ease associated with the use of the system)
3. social influence (the degree to which an individual perceives that important others believe he or she should use the new system)
4. facilitating conditions (the degree to which an individual believes that an organizational and technical infrastructure exists to support use of the system)

Reflecting on this, I believe it would have been a better choice to use the UTAUT-model, instead of the TAM-model. With the UTAUT you also look into the perceived ease of use and perceived usefulness, but you also look into the social influence and facilitating conditions. These last two aspects would have been very useful to get an idea of what people think and want from Spotify's/Soundcloud's community features and whether discovering new music is facilitated enough by the apps.

HED/UT (Hedonic Utility Scale)

Paper read: K. E. Voss, E. R. Spangenberg, and B. Grohmann. Measuring the hedonic and utilitarian dimensions of consumer attitude. Journal of Marketing Research, 40(3):310 -- 320, 2003.

With the HED/UT both the hedonic aspects, and utility and usability aspects are addressed. This could have been applied to our research question, in order to get an idea of the hedonistic differences between Soundcloud and Spotify. However, as we were interested in the usefulness and use of these to apps, TAM was a better choice.

Two other questionnaires (Selim)

Attrakdiff: focusses on the reflection of experiences, and not the actual experience. Asses hedonic and pragmatic values. Intrinsic motivation inventory: subjective assessment.

I believe that the TAM questionnaire was a better choice than these ones because it is a more objective evaluation. Also the clusters used in this method does fit our research question very well. On the other hand it might be too holistic to compare the two apps by only evaluating perceived usefulness and ease of use. By using for example the Attrakdiff method we would have a better insight on how the user feels towards the applications I believe. But then again this was not the purpose of our research.

Two other questionnaires (Arthur)

My personal choice of questionnaires as an alternative are the UTAUT² and the AttrakDiff³ questionnaire methods.

The UTAUT method (abbreviation of Unified Theory of Acceptance and Use of Technology) is a research method is mainly used in the context of computer applications to examine usage behaviour. It is used to consider four dimensions: *performance expectancy*, *effort expectancy*, *social influence* and *facilitating conditions*. It is similar to the TAM questionnaires in the sense that the usefulness of a product is considered, yet the UTAUT method adds more dimensions. A disadvantage of this is that the data generated from such a method can be all over the place, making it hard to draw a solid conclusion.

² UTAUT, <http://www.allaboutux.org/utaut>

³ AttrakDiff, <http://www.allaboutux.org/attrakdiff>

AttrakDiff is a useful, online research method that can be used to compare two products, which gives you insight in how users compare each of the products in hedonic and pragmatic parameters. This is a questionnaire that I think has great potential to be used in a context like we have, since the online functionality really speaks to me. After visiting the web-site I've noticed just how easy it is to set up a simple research, and having an online test makes gathering participants an easy task.

In hindsight, I think that we made the right decision to do our research in the format of Technology Acceptance Model questionnaires. It fit our goal well, and allowed us to solve our research question on the perceived usefulness and perceived ease of use.

3.2: Individual homework assignment analysis

For the last week of the assignment, we got the homework to make some exercises about statistical methods. This way we would become acquainted with them, and get an idea on how and when to use the different methods. This sub-chapter shows the homework of each member of the team.

Results of the statistical exercises (Selim)

1. From the calculator it becomes clear the sample size should be around 87 participant to have a power of 0.80. This was derived by choosing different sample size in order to meet this power.

2. A. These are paired samples since the concept A and B are not equal. This means that the difference is from within the designs.

B. This is a nominal scale of measurement.

C. From the test I can conclude that a sample size of at least 197 is needed to have a power of 0.80.

3. A. -

B. These are related(paired) samples.

C. $t=2,06$ degrees of freedom= 14 $p=0.058$

D. The measurement scale is ordinal

E. A T-test compares the differences between the means of two samples. It also check if the mean of the standard deviation differs from a certain value.

4. A. The P-value is 0.76

B. the data does not give me any reason to conclude that the population median differs from the hypothetical median. Plus the result is not significant at $P>0.05$

C. It compares two different samples to see if their means ranks differ.

5. A. The U-value is 398, so the distribution is approximately normal.

Results of the statistical exercises (Rosa)

1.
 - a. Related/within design
 - b. They are frequencies, so nominal
 - c. The population group should be 84 participants. **according to answers: N = 85**
These calculations are made for 2 independent study groups with a dichotomous endpoint.
2.
 - a. Related samples/within design.
 - b. Again frequencies, so nominal.
 - c. N = 31
3.
 - a. -
 - b. Within design/related.
 - c. $t = 2,06$
degrees of freedom = 14
 $p = 0,58$
 - d. A test of the null hypothesis that the difference between two responses measured on the same statistical unit has a mean value of zero.
 - e. Ordinal data
 - f. No, it is not correct. T-test is a parametric test, which is used for interval and ratio scale.
However, everyone does parametric tests on scale data. The equivalent of a t-test on non-parametric data is the Wilcoxon.
4.
 - a. W = 31
 - b. $p = 0.09894$
5.
 - a. U = 77
 - b. $p = 0.148$

3.3: Statistical analysis of questionnaire

Distribution diagrams

We started our analysis by calculating averages and making distribution diagrams. We made 2 diagrams, one for each cluster of the questionnaires. The horizontal axis represents the score which was variable from 1 to 7. The vertical axis represents the frequency which is how many times each score has been selected.

Category	Spotify	Soundcloud
Perceived usefulness (average)	4.39	4.81
Perceived ease of use (average)	4.92	4.98



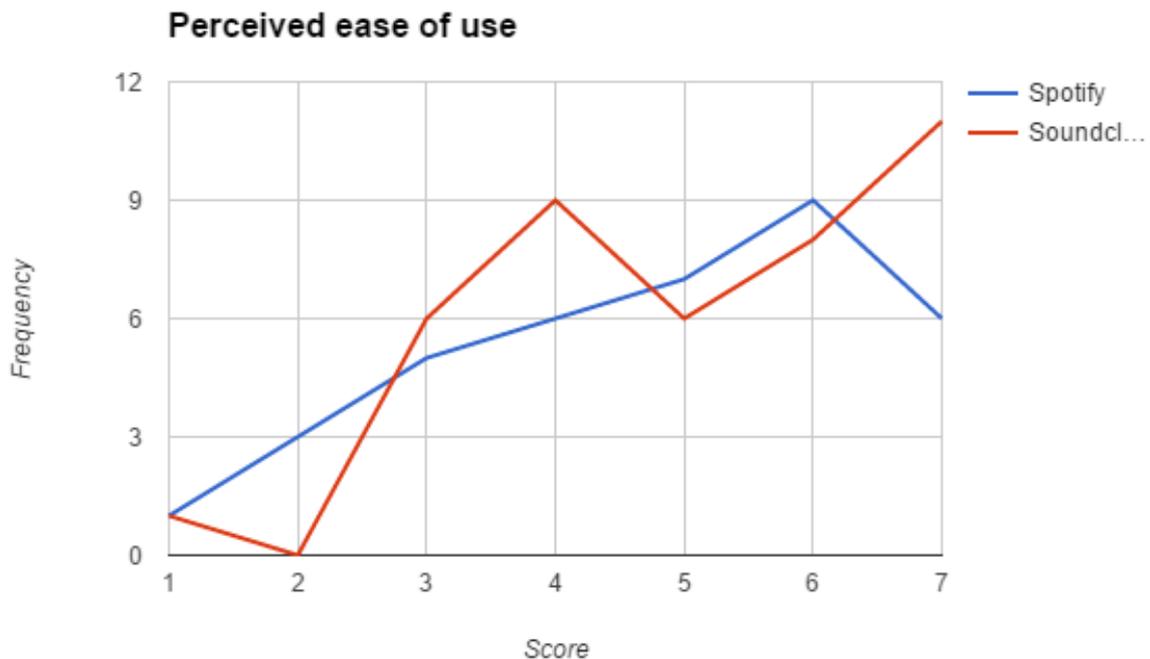
U-Test

Since our data is on a nominal measurement scale, we chose to use the Mann-Whitney test. For this test we have kept both clusters separated to get 2 p values.

Perceived Ease of use:

The p-value is 0.92034. The result is not significant at $p \leq 0.05$. Since the p-value is much bigger than 0.05 we now need to calculate how sample size should be in order to have a p-value smaller or equal to 0.05.

Sample size: from calculations we can conclude that for statistical power of 80% the sample size should be at around 1700 participants.



Perceived Usefulness:

The p-value is 0.4009. The result is not significant at $p \leq 0.05$. Since the p-value is much bigger than 0.05 we now need to calculate how sample size should be in order to have a p-value smaller or equal to 0.05.

Sample size: from calculations we can conclude that for statistical power of 80% the sample size should be at least 35.

Conclusion:

As predicted, the sample groups are not large enough to give a significant answer to the research question. Therefore, we calculated the needed sample sizes in order to have a $p \leq 0.05$ which are $n=1700$ for perceived ease of use and $n=35$ for perceived usefulness. From the means we have calculate we can conclude that Soundcloud has overall higher perceived usefulness and ease of use. The difference between the perceived usefulness between the 2 apps can be related to the fact that Spotify has 2 versions. The participants of the questionnaire were a mix of Spotify and Spotify premium users. But the results do still have a low significance because of the small sample size.

4. Personal Reflections

This chapter contains the reflections from each team member on this course. It contains the following topics: motivation for taking this course, expected learning activities, newly acquired high-level insights, overall thoughts on the course, room for improvement and the teamwork of the group.

4.1 Selim Haase

Last quartile I did the elective “UFP Basics”. After that elective I wanted to do a follow-up, so I’ve decided to do this course too. I wanted to gain knowledge of different user evaluation methods, but I also wanted to apply them in practice. Since I already did an elective on UFP, I expected that this course would go deeper in the subject matter than previous one.

I am now not only aware of the different user evaluation methods, but also know their properties and how to conduct effectively. The 3 different user evaluation methods that were presented to me can all have their space during different phases of the design process. Having to work in groups was a pleasant experience because we could discuss the subject of matter from different perspectives, which I found valuable. I believe that the teamwork went well since all of the members were motivated and eager to learn. By keeping contact regularly and discussing the theory we were able to use the methods and techniques during homework. Doing weekly presentation has helped to exercise and develop some professional skills such as presenting.

Where “UFP basics” stopped at user evaluating, this elective went a step further by validating the results through different methods of analysis. I’ve learned to select an appropriate analysis method based on the type of data generated. Assessing the quality of the data was also new for me. I became aware that by assessing and reflecting on the data, room for improvement can be created to get more reliable results. Even though I had to conduct and analyse my own research as an exercise, I do believe that by implementing on my own work I can make this process more refined and less time consuming. For future work and my projects, these methods will help me validate and evaluate my work through user involvement in different stages of the process.

4.2 Arthur Geel

My main motivation in signing up for this elective was because I wanted to get a better understanding on how to properly conduct user tests, but also on how to analyze the results to get the most out of them. Before this course I was very inexperienced with user testing, it was something I would do at the end of a design process, mainly because I was expected to have user test data for the final report. In retrospect, I think my passive stance towards user testing was because I did not fully recognize what user testing has to offer in various stages of design processes.

During the lectures I learnt all about users in design processes. We were given reasons why we should include users in design processes in varying stages, as this can help the designer to understand the user and the user in a specific context, the goals a user has and the needs a user has. We were also given information about how to involve the users in a process. To me, three methods stood out: *contextual inquiries*, *diary studies* and *cultural probes*. Every one of these has their own advantages and disadvantages, though my personal favourite was the cultural probe.

Cultural probes allow you to swiftly gain overall insights in the needs from a user's perspective, which makes them a powerful tool to have at the start of a design process. Later on in the course I learnt about the different types of research, *qualitative* and *quantitative research*, and for what reason they should be used, for *generation* or for *evaluation*. In the workgroup meetings I was able to discuss these concepts with my peers, and later on I was able to sufficiently implement this knowledge by setting up various research iterations.

The second half of this elective was focused on analysing the gathered data on multiple dimensions: qualitative/quantitative-wise and subjective/objective-wise. We were taught ways to look back at the research and examine if it was done properly by looking at the protocol, the sample size and on significance. I was introduced to statistical tests such as Wilcoxon's tests, t-tests, u-tests and numerous others. Therefore, I am now able to use these to verify the validity of the research I have conducted.

In retrospect, I think the build-up of this elective was well planned, and the contents were up to my expectations. I enjoyed the balance this course had with lectures and subsequent workgroup sessions, as they motivated me to apply the newly learnt theory in practice. This course allowed me to gain understanding of the importance and the process of user testing and analysing the data. I now have the tools to incorporate in-depth user testing in my overall process, which will help me increase the quality of my work. I consider this course to be a useful asset to my overall development as a designer, which is why I would recommend it to fellow designers who want to get an introduction in conducting research through user tests.

4.3 Rosa van Koningsbruggen

Before this assignment, I did not have real knowledge on how to properly conduct user tests or experience with it. Besides that, my B22 semester has to be a research project. For these reasons I decided to follow the assignment Designing for the User Experience. With this assignment I wanted to learn how to include the user into the beginning of the design process and how to validate your design at the end of the process. Especially the part of including a user during the design process seemed appealing to me, as this knowledge would be valuable for my research project.

During the first half of the assignment, I was introduced to qualitative methods that you could use during the begin of the design process. By reading the papers on the different methods and having weekly presentations, I have gotten an idea of when to use what method. As a group we chose to execute the cultural probe method, as this would allow use to get a personal insight in people's music listening experience. From working with this method, I have experienced that it is a valuable way to gain data on what the intended user wants, thinks, and needs. Assumptions you have as a designer beforehand can be contradicted or proofed to be true. Either way, I believe that including the user in such a way will improve your design.

The second half of the assignment was on quantitative data and validating. During this part of the assignment I have learned how to select the appropriate questionnaire and how to analyse this data. This turned out to be much harder and rigid than I thought. I had never given it a thought that questionnaires themselves have to be validated and that you can calculate the number of participants you need to have in order to obtain significant results. This knowledge will be useful for all my future questionnaires and will give me a certain perspective on when you can trust data.

5. Appendices

Appendix A: Task division

Task division user tests

User Test	Participants	Recruited by	Made by
Pilot	4	Arthur, Selim, and Rosa	Arthur, Selim, and Rosa
Cultural Probe	9	Arthur, Selim, and Rosa	Arthur, Selim, and Rosa
TAM-method questionnaire	18	Arthur, Selim, and Rosa	Arthur, Selim, and Rosa

Task division report

Chapter	Done by
1	Arthur Geel
2	Rosa van Koningsbruggen
3	Selim Haase

Appendix B: Raw Questionnaire data

Spotify

X-axis = score

Y-axis = question number

	1	2	3	4	5	6	7
1		1	1		3	4	
2		1	1	1	3	2	1
3			1	2	2	4	
4			2	1	3	3	
5	1	1	2	2	2		1
6		2	3	1	1	1	1
7	3	1	1	1	1	2	
8				6	2	1	
9				1	2	4	3
10		1	1		2	3	2
11	3	1		2	3		
12			1	5	2	1	
13				3	2	3	
14		1	2	3	1	2	

Soundcloud

X-axis = score

Y-axis = question number

	1	2	3	4	5	6	7
1	0	0	1	1	3	2	2
2	0	0	2	4	2	0	1
3	0	2	1	0	1	3	2
4	0	0	1	0	2	2	2
5	1	0	1	2	1	3	1
6	0	0	2	3	2	0	2
7	2	1	2	2	2	0	0
8	0	1	2	2	3	1	1
9	0	0	2	1	1	2	2
10	0	0	1	1	1	1	4
11	3	1	1	1	0	1	
12	0	0	1	2	3	2	1
13	0	2	0	2	1	1	4
14	1	0	0	2	1	2	2

Appendix C: User tests and data

The raw data of the user tests have been uploaded to Google Drive. This was done to ensure that this report would not become too big in file size, since they are quite extensive.

Cultural probe: Raw data:

<https://drive.google.com/open?id=0B7CKnggC-UqIX3Z6VHIYMGZ6Tzg>

The same goes for the results of the StoryPly session in week 4. To save bandwidth, we have uploaded the raw data to Google Drive.

StoryPly: Raw data

<https://drive.google.com/open?id=0B7CKnggC-UqIM2RId2UzcGJEOUk>

Appendix D: Summaries

Summaries Designing for the UX
Rosa van Koningsbruggen

Elementary Quantitative Data Analysis

quantitative data analysis = statistical techniques used to describe and analyse variation in quantitative measures/numbers to discover and describe patterns in your data.

statistic = a numerical description of a population, usually based on a sample population.

Used for:

- describing the results of measuring single variables/constructing and evaluating multi-item scales.
 - describing the association among variables and to control for other variables.
- These statistics are called *descriptive statistics*, as they describe the distribution of and relationship among variables.

inferential statistics = to estimate the degree of confidence that can be placed in generalizations from a sample population from which the sample was selected.

secondary data analysis = analysis of data collected by someone other than the researcher or the researcher's assistants.

How to analyse?

- I. Prepare data in a format that is suitable for computer entry. Also, conduct a *data cleaning* = the process of checking data for errors after the data have been entered in a computer file.
- II. Discover the variation in each variable of interest. The two most popular formats used for this are graphs and frequency distributions. These create a distribution shape, which can have 3 important features:
 1. *central tendency* = the most common value/the value around which cases tend to centre. It's usually summarized with one of three statistics:
 - § *mode* = the most frequent value in a distribution. It is not as frequently used as the other 2, as it can easily give a misleading impression. Furthermore, there occurs a problem when a distribution is *bimodal* (has 2 categories with a roughly equal number of cases and clearly more cases than the other categories), instead of *unimodal*.
 - § *median* = the position average, or the point that divides the distribution in half.
 - § *mean* = the arithmetic average.
 2. *variability* = the extent to which cases are spread out through the distribution or clustered around just one value

3. *skewness* = the extent to which clusters more at one or the other end of the distribution of a quantitative variable rather than in a symmetric pattern around its centre. Positive skew when most cases are tapering off in the positive, otherwise, the skew is negative.

III. Variation

§ *range* = highest value – lowest value + 1

An disadvantage of this method is that it can be drastically altered by just 1 extreme value = *outlier*.

§ *interquartile range* = the range in a distribution between the end of the 1st and the beginning of the 3rd quartile. *quartile* = the points in a distribution corresponding to the first 25% of the cases, the first 50%, and the first 75%.

§ *variance* = a statistic that measure the variability of a distribution as the average squared deviation of each case from the mean.

§ *standard deviation* = the square root of the average squared deviation of each case from the mean.

cross-tabulation = displays the distribution if one variable within each category of another variable (it shows 2 variables at the same time and can even be used for more variables). It is often shown in a crosstab-table.

gamma = a measure of association that is sometimes used in cross tabular analysis.

chi-square = an inferential statistic used to test hypotheses about relationships between 2 or more variables in a cross tabulation.

It can also be used for 3 variables, where the third one is, for example, an *extraneous variable* = a variable that influences both the independent and dependant variables so as to create a spurious association between them that disappears when the extraneous variable is controlled.

elaboration analysis = the process of introducing a third variable into an analysis to better understand the bivariate relationship under consideration.

statistical significance = the mathematical likelihood that an association is not due to chance, judged by a criterion the analyst sets.

Using Graphs and Charts to Illustrate Quantitative Data

Selecting how to present your data:

- what do you want to present?
- which variables do you want to include?
- should those variables be expressed as frequencies, percentages, or categories?

When you are selecting a graph/chart, ask yourself what kind of data are you working with?

1. categorical data: grouped into non-overlapping categories, so bar graphs, line graphs, and pie charts;

2. continuous data: are measured on a scale or continuum, for example a histogram.

5 Ways to Avoid Being Fooled by Statistics

Processes from data to statistics:

- data collection
- data entry
- data analysis
- data reporting
- data visualisation

Check data for correctness:

1. Use maths and common sense;
2. Look for the source and check the authority of the source;
3. Are the statistics biased or statistically insignificant? (this is for the phases data collection and data analysis). The following are frequently seen:
 - § *sampling bias*
 - § *statistically insignificant*
4. are the statistics skewed purposely or misinterpreted?
5. fully utilize your resources.