User Experience Evaluation of a Streaming Music System
An Explorative Study of Spotify’s Desktop Client

Scientific Theory and Methodology
Examination paper, 7,5 ECTS credits

Date: 2016-10-24
Supervisor: Odd Fredriksson
Abstract
Music is a domain that is subjectively driven in terms of people having different tastes and preferences among various genres. In the later years of music there has been development in digitalization because of how people listen, store, and distribute it. The music world has moved from CDs and has now found its place – among other areas – in a streaming domain where endless amount of musical content can be found across multiple devices. The academic works within the music system realm has traditionally had its focus on system-centered approaches rather than user-centered approaches. It is only in later years that it has become apparent there needs to be more studies that covers and focuses on people’s music information needs and behaviors in terms of user experience. One of the many commercial music services is Spotify which is a streaming platform for music. The purpose of this study is to identify and describe Spotify’s desktop client’s user experience positives and flaws. This study outlines important user experience aspects (usability, utility, and aesthetics). To evaluate the system an online questionnaire was created and also usability tests accompanied with interviews were conducted. The main conclusions drawn from this study are that the desktop client: cover the needs of most users; basic functionality (e.g. searching for and playing songs) is well communicated and explained whereas more advanced functionality is not, and; users consider the desktop client to be effective and efficient to use.

Keywords: User experience, streaming music systems, Spotify, usability, utility, aesthetics.
# Contents

1. Problem background ........................................................................................................... 4  
   1.1 Purpose ................................................................................................................................ 4  
   1.2 Empirical case ...................................................................................................................... 4  
2. Main theoretical perspective ................................................................................................... 5  
   2.1 User experience ..................................................................................................................... 5  
   2.1.1 Aspects of user experience .............................................................................................. 5  
   2.2 Evaluating user experience in music systems ....................................................................... 6  
   2.2.1 Usability .......................................................................................................................... 6  
   2.2.2 Utility ............................................................................................................................. 7  
   2.2.3 Aesthetics ......................................................................................................................... 8  
   2.3 Conceptual framework ......................................................................................................... 9  
3. Method ...................................................................................................................................... 10  
   3.1 Online questionnaire ............................................................................................................ 10  
   3.2 Usability tests ....................................................................................................................... 11  
      3.2.1 Pilot test .......................................................................................................................... 12  
      3.2.2 Background questionnaire ............................................................................................... 12  
      3.2.3 Usability task .................................................................................................................. 13  
      3.2.4 Unstructured interview ................................................................................................... 13  
4. Results and data analysis ......................................................................................................... 14  
   4.1 Usability .................................................................................................................................. 14  
      4.1.1 Overall satisfaction .......................................................................................................... 14  
      4.1.2 Performance ..................................................................................................................... 15  
      4.1.3 Learnability ...................................................................................................................... 16  
      4.1.4 Affordance ...................................................................................................................... 16  
      4.1.5 Feedback ........................................................................................................................ 17  
   4.2 Utility ...................................................................................................................................... 18  
      4.2.1 System features matching user needs .............................................................................. 18  
   4.3 Aesthetics .............................................................................................................................. 19  
      4.3.1 Attractiveness ................................................................................................................. 19  
      4.3.2 Clearness ........................................................................................................................ 19  
      4.3.3 Creativity ....................................................................................................................... 20  
   4.4 Analysis of research design ................................................................................................... 21  
5. Conclusions ............................................................................................................................... 21  
   5.1 Future studies ....................................................................................................................... 22  
Appendix .................................................................................................................................... 26  
   Appendix 1. Online questionnaire in Google Forms. ................................................................ 26  
   Appendix 2. Test plan ............................................................................................................... 27  
   Appendix 3. Summarized text from the unstructured interviews structured. .......................... 30
1. Problem background

Music is a subjectively driven domain in a sense that people typically often have a taste and preference across various genres; people normally prefer to listen to different kinds of music. "Music that holds no interest for a given subject very frequently ‘sounds the same’" (Berenzweig et al. 2004). During the later years there has been advancement in digitization of music – how people listen, store, and distribute it. It was not long ago people listened to CDs (or potentially cassettes) and now we situate ourselves in a world where we can stream an endless amount of musical content by the click of a button through various types of channels (smartphones, laptops, etc.).

One of many commercial music services is Spotify (2016) which is a streaming platform for music. Spotify has throughout the years expanded their functionality by, e.g., creating personalized playlist and venturing the world of displaying videos (by smartphone) to its audience (Welch 2015). According to Internetstiftelsen i Sverige (2011), 57% of the Swedish population listens to music on the Internet and 37% streams music through Spotify every month; 9 out of 10 in the years 16-25 listens to music on Spotify. This shows the big change in presentation of music for consumers. As the technology has advanced throughout the years new issues arrived onto the scene of music streaming services, such as: privacy concerns, compatibility among devices, and limitation of storage (Lee et al. 2015a). The science of retrieving information from music is called Music information retrieval (MIR) which is a small research field (Lee & Price 2015). A big portion of the work within MIR has traditionally focused on system-centered approaches rather than user-centered approaches; something that has been pointed out in numerous works (e.g. Hu et al. 2015; Schedl et al. 2013). It’s only in later years that user experience is being evaluated as well (Lee & Price 2015); by focusing on people’s music information needs and behaviors.

As the research area is making progress there is still a lot of ground to cover; especially in the context of the front-end of the systems. Other studies (Lee & Price 2015; Lee et al. 2015b) have focus on more than one commercial service. Lee et al. (2015b) state that there should be focus on evaluation rather than competition: “If we treat this as an opportunity to evaluate the systems in order to improve the design of all participating systems rather than being able to claim one system is better than the other, this issue will naturally dissolve”. A few examples of usability and user experience issues were showcased in a study conducted by Lee and Price (2015) where services such as Spotify, Pandora, and other music services, were widely criticized by the study’s participants. Therefore, more studies need to be conducted regarding the various music services with a use-centered perspective; as there are plenty of different services, different functionalities, and different platforms (i.e. smartphones, desktops, operating systems and other mixes) to examine.

1.1 Purpose

This study is solely focused on the commercial music service Spotify and its desktop client. The purpose is to identify and describe its user experience positives and flaws.

1.2 Empirical case

There are two types of strategies of conducting research: quantitative and qualitative (Bryman 2012). Qualitative research is a research strategy that focuses on in-depth data acquired through, e.g., interviews and verbal analysis methods (Patel & Davidson 2011). Quantitative research on the other hand infers measurements in data collection and statistical processing (Patel & Davidson 2011). Rohrer (2014) states that the researcher, in usability studies, observes how people use systems to meet their needs. By observing it gives the researcher the option to ask questions, probe on behaviors, or adjust the test protocol to better meet its objectives; often in these scenarios the data is not mathematical. Rohrer (2014) concludes that quantitative methods help to see what issues have the biggest impact. This study will have a combination of qualitative and quantitative research methods. The qualitative method for this study will be usability tests with the purpose of observing participants with accompanied interview questions to uncover and get in-depth information regarding the participants’ perceptions of the system. As for the quantitative method an online questionnaire will be conducted to find out how the consumers experience the system.

It is crucial in the testing process of a system that the selection of participants matches the intended end user (Dumas & Redish 1999; Rubin & Chisnell 2008). The participants were approached by the researcher knowing about their usage of Spotify as it is important to define the criteria for the user
group of a system (Rubin & Chisnell 2008). Therefore, it was important to create background questions that concerned experience characteristics to receive answers from the right users.

Research – that involves users - needs to take ethical considerations into account. To develop knowledge it is important that the researcher understands the responsibilities of how to act concerning the participants (Patel & Davidson 2011). Bryman (2012) discusses the importance of treating people on whom we conduct research; e.g. addressing not harming participants and if there is lack of informed consent or not. There are multiple requirements on what to inform the participants of (Vetenskapsrådet 2002):

- **Information**: entails that the researcher should inform participants of the current research’s purpose.
- **Consent**: entails that the participants have the right to decide for themselves if they want to participate or not.
- **Confidentiality**: entails that all participants should be granted maximum confidentiality and the personal disclosures should be stored in a way that unauthorized people cannot access the information.
- **Use**: entails that collected data about participants can be used for the research’s purpose.

By these remarks this study made it clear to the respondents in the online questionnaire and the participants involved in the usability tests of: the research purpose, the involvement of the participants and their rights, the anonymity, and where the information was going and how it was used.

### 2. Main theoretical perspective

This chapter will outline the various factors concerning this study and how it is adopted to the research field of music services.

#### 2.1 User experience

Norman (1998) initially devised the term *user experience*. The ISO Draft International Standard (2008) defines user experience as: “A person’s perceptions and responses that result from the use or anticipated use of a product, system or service”. Norman (2016) explains that the term has been misused by people who design, e.g., websites or application when in reality the concept of user experience is broader than that. Hassenzahl and Tractinsky (2006) believe that it is not accidental that the term has become a buzzword because of the many interactive products into our daily lives.

Norman (2016) states that user experience is everything that touches upon the user’s experience with the product (or even explaining to someone else about the product). The user experience pioneer said that the term was invented because the concepts of human interface and usability were too narrow (Merholz 2007). Other literature helps us further understand the term user experience (Garrett 2010:7):

> When most people think about product design … they often think of it in terms of aesthetic appeal: a well-designed product is one that looks good to the eye and feels good to the touch. (The senses of smell and taste don’t come into play for most products. Sound is often overlooked but can be an important part of the aesthetic appeal of a product.) … Another common way people think about product design is in functional terms: A well-designed product is one that does what it promises to do. And a badly designed product is one that somehow doesn’t … products might look great and work well functionally, but designing products with the user experience as an explicit outcome means looking beyond the functional or aesthetic.

Norman and Nielsen (n.d.), of Nielsen Norman Group, explain that it is important to distinguish the total user experience from the user interface; even though the interface is an important part of the design. Mahlke (2005) states that user experience can be used as an umbrella term to compile all the essential aspects of interaction from the user’s perspective. A widely presented illustration, from nnGroup, of user experience and its subcategories is often referred to when distinguishing user experience from usability (Neospot 2010). The idea is that for a product to achieve required user experience the following subcategories are needed: **usability** which entails I am able to use the product easily; **utility** which entails it is useful to me and it meets my needs; **desirability** which entails I like the way the product looks and feels, and; my overall feeling about the brand/product is good.

#### 2.1.1 Aspects of user experience

Norman and Nielsen (n.d.) states that user experience encompasses all aspects of the end-user’s interaction with the company, its services, and its products. But what are then the aspects? In multiple
studies the concept of user experience are composed by multiple parts; these are often referred to as components, elements or aspects. Garrett (2010) defines various elements of user experience: the surface plane, the skeleton plane, the structure plane, the scope plane and strategy plane. These different aspects work as a blueprint for achieving user experience for a product (physical or virtual). This is a structural view that creates an overview of what to think of when creating a product. Forlizzi and Battarbee (2004) believe that user experience is associated with a variety of meanings as usability, beauty, hedonic, and affective. Hassenzahl and Tractinsky (2006) extend the discussion of user experience by mentioning three different perspectives with accompanied parts: beyond the instrumental (holistic, aesthetic, and hedonic); emotion and affect (subjective, positive, antecedents and consequences), and; the experiential (dynamic, complex, unique, situated, and temporally-bounded). Another study shows similar aspects as it highlights the importance of satisfaction and loyalty through the utility, ease of use and pleasure provided in the interaction with a product (Kujala et al. 2011). However, the aforementioned aspects do not specifically mention any type of system but try to generalize the concept which in itself is contextually essential. It is therefore important to look at aspects of user experiences in studies concerning music services.

Music services, depending on the system, are often called recommender systems in academic contexts. Recommender systems are systems which allow users to find valuable items depending on their interests (Miller et al. 2004) – items are probably derived from history results where the user has been looking for an object. There are studies (Schafer et al. 1999; Chen & Pu 2014) of evaluating user experience that revolves around the way people feel about the product and their satisfaction. However, there is no standardized rule set for examining the user experience in music services (recommender system or not).

2.2 Evaluating user experience in music systems

Hu et al. (2015) explain that there are no evaluation frameworks in music services focusing on complete systems and user experience. Therefore, Lee et al. (2015b) designed, implemented, and tested a holistic user experience evaluation platform as a step toward a comprehensive evaluation framework. As pointed out by Hu et al. (2015) the framework emphasizes the evaluation of user experiences of the entirety of the system rather than only evaluating search results. In the following subsections of this chapter different aspects will be covered for evaluating the user experience of music systems.

2.2.1 Usability

International Standard Organization (1998) defines usability as the “extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use”. Abran et al. (2003) explain that effectiveness can be interrogated as: “How well do the users achieve their goals using the system?”, and efficiency as: “What resources are consumed in order to achieve their goals?”. The last attribute of usability, according to International Standard Organization (1998), is satisfaction which is closely related to user experience as it is an overall assessment of the user’s evaluation. Satisfaction can be interrogated as: “How do the users feel about their use of the system?” (Abran et al. 2003). Nielsen (2012a) defines usability as a quality attribute that detones how easy a user interface is to use. Consequently, Nielsen (2012a) mention five components that defines usability: (1) learnability: how easy is it for users to accomplish basic tasks the first time they encounter the design?; (2) efficiency: once users have learned the design, how quickly can they perform tasks?; (3) memorability: when users return to the design after a period of not using it, how easily can they reestablish proficient usage of the system (i.e. effective usage)?; (4) errors: how many errors do users make, how severe are these errors, and how easily can they recover from the errors?, and; (5) satisfaction: how pleasant is it to use the design?.

Hu et al. (2015) collect usability principles from Nielsen (1995) and Nielsen and Molich (1990) which are widely used and well-accepted usability heuristics can be adoptable to user experience. These heuristics are: visibility of system status; match between system and the real world; user control and freedom; consistency and standards; error prevention; recognition rather than recall; flexibility and efficiency of use; aesthetic and minimalist design; help users recognize, diagnose, and recover from
errors, and lastly; help and documentation. Hu et al. (2015) compile these heuristics into four different factors to examine the usability of a music system:

- **Learnability:** how easy was it to figure out how to use the system? Learnability entails ease of use and help as it should be simple, intuitive, and user-friendly (Lee et al. 2015b).
- **Robustness:** how good is the system’s ability to warn you when you’re about to make a mistake, allow you to recover, or retrace your step?
- **Affordance:** how well does the system allow you to perform what you want to do? Lee et al. (2015b) explain that affordance is about giving the user access, presenting simple functions and being able to browse/search for specific songs/artists/albums etc.
- **Feedback:** How well does the system communicate what’s going on? This factor entails ‘clarity’ as it is about clear and intuitiveness functions and labels (Lee et al. 2015b).

However, one of the key outlines made by Lee et al. (2015b) – in a study where these principles were used – made the point that the criterion of robustness was difficult to evaluate because of limited time in a study. This is most likely because of trying to trigger errors and letting people recover while using a system. Therefore, Lee et al. (2015b) propose to rethink the criterion.

Abran et al. (2003) present a usability model of two groups of researchers (Dix et al. 1993; Nielsen 1994) with the same set of usability characteristics; concerning effectiveness, efficiency, satisfaction, and learnability. Dix et al. (1993) as referenced by Abran et al. (2003:332):

- **Effectiveness:** percentage of tasks accomplished, ratio on failure of handling, and percentage of tasks achieved per unit of time.
- **Efficiency:** number of good and bad characteristics recalled by users, time to complete a task, time spent on errors, error’s percentage, documentation or help’s use frequency, repetitions’ number of failed commands, and number of available commands not called upon.
- **Satisfaction:** percentage of users’ favorable and unfavorable comments, number of times that user expresses his frustration, and rating scale for users’ satisfaction with functions and characteristics.
- **Learnability:** time to learn.

Nielsen (1994) as referenced by Abran et al. (2003:332):

- **Effectiveness:** percentage of tasks accomplished.
- **Efficiency:** time to complete a task, time spent on errors, error’s percentage, documentation or help’s use frequency, and repetitions’ number of failed commands.
- **Satisfaction:** number of times that user expresses his frustration, and rating scale for users’ satisfaction with functions and characteristics.
- **Learnability:** time to learn.

Abran et al. (2003) specify that it is necessary to measure the performance and satisfaction of users working with the product to verify that an appropriate level of usability is achieved. Efficiency and effectiveness are performance criteria (Rubin & Chisnell 2008) and easily obtainable through usability testing. Effectiveness can be measured by how many answers are correct (Jeng 2005); in a usability test this could be observed by the number of completed tasks by the users. An example of measuring efficiency could be by the amount of time it takes to complete a task and/or how many clicks/steps/movements the user needs to make (Jeng 2015). Lee et al. (2015b) explain that performance is one of the factors that need to be considered in future user experience evaluations. However, satisfaction is a subjective factor that can differ depending on the user and can be examined by using a post-test questionnaire using a Likert scale (Jeng 2015; Lee et al. 2015). A criterion in the framework by Hu et al. (2015) uses the label of ‘overall satisfaction’ to measure music systems. The term overall satisfaction originally comes from Pu et al. (2011) in the context of investigating recommender systems.

### 2.2.2 Utility

Lee et al. (2015b) recommend utility being one of the potential criteria to involve in user experience studies. Nielsen (2012a) defines the concept of utility as something that provides the features you need; i.e. Does it (e.g. the system) do what the users need? Utility, together with usability, becomes something that is useful (Nielsen 2012a). Garrett (2010) explains that user needs is an important part of user experience. Utility is mentioned in multiple studies within recommender systems and/or in

Tricot (2007) states that utility is synonymous with relevance or efficacy and that the concept answer an interrogatory state of: “Does the external representation allow the learner to reach his/her leaning goal?”. Lee et al. (2015b), on the other hand, only explained the factor of utility as usefulness. Nielsen’s (2012) definition of utility, on the other hand, does not mention the term usefulness until the term utility is combined with usability. Shackel (1991) state that utility is of critical relevance as it needs to answer the question “will it do what is needed functionally?”. Nielsen (2012a) explain that the factors of usability and utility are equally important and it matters a little that something is easy to use if it is not what you want to do.

2.2.3 Aesthetics
Aesthetic is a factor that should be involved in future user experience evaluation studies (Lee et al. 2015b). Garrett (2010) explain that functionality make sure it triggers the correct action on a device whereas the aesthetic design make sure that the button is an appealing shape and texture. Both of these aspects are important to the user experience. Moshagen and Thielsch (2010) makes the point that it has become more apparent that user needs go beyond just being useful (usability and utility) and takes on user experience which includes emotions and visual aesthetics. Aesthetic data visualization are also shown to affect usability (Cawthon & Moere 2007); showing a connection between the factors in the context of user experience.

What is often referred to in user experience studies are the aesthetics of a system (Preece et al. 2002; Mahlke 2007; Cawthon & Moere 2007; Lee et al. 2015b); aesthetic in itself is often referred to something general and abstract such as: attractive, beautiful, and/or visually pleasing. Lee et al. (2015) defines the category – in their framework – to when the user specifically talks about the visual appeal of the interface. One might think it would be easy to think of aesthetics aspects in computer-related subjects as, e.g., colors, shapes, and other visual objects. However, most of the user experience work – and/or other related user interface studies – has not evolved into specific contributing indicators of the widely used term aesthetic. As showcased aesthetic is a difficult topic to define into extensive qualities to examine a system. When the term user experience first became popular within academic work the technology-oriented field was reluctant to consider hedonic qualities of creating an interactive user experience design (Butler 1996). Lavie and Tractinsky (2004) specifically mention the correlation between aesthetic and usability. Hartmann et al. (2008) suggest, from the findings in their study, that aesthetic could be an important determinant of user satisfaction and thereby overcoming usability issues and positively influencing content. Consequently, Lavie and Tractinsky (2004) state that the factor of aesthetic serves a purpose in the visual representation of websites but also the entire user experience. Lavie and Tractinsky’s study (2004) is one of the more known aesthetic studies in as much as they have identified qualities such as: clean, clear, pleasant, and symmetrical. These qualities are commonly called classical aesthetics as they are emphasizing on order and clean design. Lavie and Tractinsky (2004) also defined expressive aesthetics which includes attributes such as: original, sophisticated, and creative. Expressive aesthetic refer to creative and unconventional beauty. Park et al. (2004) presents 13 critical factors of aesthetics: bright, tense, strong, static, deluxe, popular, adorable, colorful, simple, classical, futuristic, mystic, and hopeful.
2.3 Conceptual framework

To reach a positive user experience in music the following factors have been collected within music related academic works: usability (overall satisfaction, performance, learnability, affordance, and feedback), utility, and aesthetics. By evaluating the different factors and their accompanied we can determine the flaws of music systems.

The usability variable consists of five different attributes where two of the attributes, satisfaction and performance, make up for the definition set by the International Standard Organization (1998). However, as presented in Figure 1, the wording of satisfaction is changed to overall satisfaction to show consistency in music system studies as Pu et al. (2011) explain that it gives the users an opportunity to express their preferences and opinions about the system at hand in a direct way. Rubin and Chisnell (2008) explain that the attribute of performance is the categorization of effectiveness and efficiency. Effectiveness is summarized as percentage of tasks accomplished, ratio on failure of handling, and percentage of tasks achieved per unit of time (Abran et al. 2003). Efficiency is summarized as time to complete a task, time spent on errors, and error’s percentage. The remaining usability attributes – learnability, affordance, and feedback – are devised from Hu et al.’s (2015) user experience framework. Learnability is the times it takes to learn (Nielsen 1994) and entails how easy it is to learn how to use the system (Lee et al. 2015b), affordance implies giving the user access (Lee et al. 2015b), and feedback entails communication.

Both the factors of utility and aesthetics are recommended to use in future user experience studies by Lee et al. (2015b). The factor of aesthetics is often categorized as a hedonic user experience attribute as it points towards beauty and attractive. An indicator commonly referred to in user experience studies is attractiveness (Lavie & Tractinsky 2004; Lee et al. 2015b). An occurring indicator of aesthetics mentioned is clearness (Nielsen 1995; Lavie & Tractinsky 2004). Lastly, creativity (or sometimes innovative) design is an aspect of aesthetics that is usually mentioned in user experience or usability studies (Lavie & Tracinsky 2004; Laugwitz et al. 2008; Pärnänen 2016).

The factor of utility is defined as the system providing the user with a needed feature (Nielsen 2012a). The factor of utility has been defined in similar ways – as presented in 2.2.2 Utility – but there still is some inconsistency of how it is used. In an attempt to easier explain utility’s purpose in this
study the indicator of ‘System features matching user needs’ was created. Figure 1 display three factors (usability, aesthetics, and utility) leading to the outcome of a positive user experience. A large portion of this presented framework concerns usability as it is a central part in user experience and mostly used in music system-related studies. Laugwitz et al. (2008) explain that usability criteria do not cover all aspects relevant for user experience as there are qualities such as hedonic parts as well.

3. Method

This study involved Spotify’s desktop client (version 1.0.38.171). Two major types of methods were used to collect data in this study. An online questionnaire was created but also complimented by an approach – closer to users – called usability testing. The usability tests were carried out using a Windows 10 operating system but no differences were to be found in terms of the tasks (and the major user interfaces) carried out by the participants (see 3.2 Usability tests) on Mac operating systems.

Spotify has various types of consumers (i.e. the users of the system): premium paying users, non-paying users, users who prefer to listen on mobile phones, users who prefer to listen on desktops, users in various ages and genders, etc. However, it is seemingly improbable to get proper demographics of Spotify’s population. According to Internetstiftelsen i Sverige (2011), 57% of the Swedish population listens to music on the Internet and 37% streams music through Spotify every month; 9 out of 10 in the years 16-25 listens to music on Spotify. CivicScience (2015) collected data – without Spotify’s participation in the U.S. – and stated that 61% of Spotify users are 29 years old and younger. Stated by Statistiska centralbyrån (2015) almost 65% of the Swedish population have used services to watch movies or listen to music on the Internet (two examples mentioned are Spotify or Netflix); approximately 90% of them are between the ages of 16-34 years old. It is evidential that all ages uses streaming services (CivicScience 2015; Internetstiftelsen i Sverige 2011; Statistiska centralbyrån 2015). However, there are still a lot of people using streaming services in the years of 35-85 but the decline in people are significant between the years of 75-85 (Statistiska centralbyrån 2015). Also, there is no way of knowing specifically what streaming services are covered and what ages utilize what service.

Even though the aforementioned statistics have some consistency there are still no official statistics presented by Spotify. It will be improbable to match the population of Spotify’s client – even though the researcher would like to – and therefore this study’s participants (in the usability tests) and respondents (in the online questionnaire) will be chosen by the researcher based on knowing of the participants Spotify usage. Rubin and Chisnell (2008) explain it is crucial that the testing process of a system must match the intended end user. Therefore, the researcher of this study chose to only ask background questions to have the appropriate people participating.

3.1 Online questionnaire

Laugwitz et al. (2008) explain that questionnaires are a commonly use tool for quality and usability within user-driven assessment. Consequently, Laugwitz et al. (2008) state that questionnaires are an efficient quantitative instrument to measure product features. Walsh and Nurkka (2012) argue that user experience web-surveys have potential to gather user data as technology services are getting accessible for wider user groups. Bryman (2012) states that an online questionnaire is completed with fewer unanswered questions than other types of questionnaires resulting in less missing data; such as postal questionnaires. Laugwitz et al. (2008) state that both ‘soft’ criteria (such as hedonic attributes such as aesthetics) and ‘hard’ criteria (usability attributes) can be measured using a questionnaire as it will be easy to apply, reliable and a valid measure because it is a complement data from other evaluation methods with subjective quality ratings. The only attribute that was shown to be a weakness in Laugwitz et al.’s (2008) study was the factor of ‘dependability’. Chen and Pu (2014) conducted an online survey questionnaire which concerned subjective user experience perception of their interfaces to complement their usability tests. Petrie and Bevan (2009) state that satisfaction and user experience is a part of the evaluation of electronic systems and can be assessed in a variety of ways; the simplest way is with rating scales and questionnaires. Also, Rubin and Chisnell (2008) explain that subjective data that measure a person’s feelings or opinions is obtainable in online questionnaires.

The questionnaire for this study was built in a way similar to the one show-cased in Lee et al.’s (2015b) user experience framework using interrogating questions concerning the different user
experience indicators presented in Figure 1; thus making it possible for users to answer using a Likert scale.

- **Overall satisfaction:** How would you rate your overall satisfaction with the system?
- **Performance:** How well does the system perform in terms of:
  - allowing you to complete tasks correctly?
  - allowing you to complete tasks time-efficiently?
- **Learnability:** How easy was it to figure out how to use the system?
- **Affordance:** How well does the system allow you to perform what you want to do?
- **Feedback:** How well does the system communicate what is going on?
- **Utility:** How well do the system’s features cover your needs?
- **Attractive:** How attractive is the system to look at?
- **Clear:** How visually clear is the system?
- **Creative:** How visually creative is the system?

The attributes of learnability, affordance, feedback, and overall satisfaction with their accompanying questions was part of Lee et al.’s (2015b) study. Therefore, the same instances of questions were used in this study because of displaying consistency within a growing field. Aesthetic was a factor that was mentioned by Lee et al.’s (2015b) study and was classified as “how appealing the visual interface was”. However, to make the factor of aesthetics less abstract three indicators (attractiveness, clearness, and creativity) were derived from multiple works (Nielsen 1995; Lavie & Tractinsky 2004; Laugwitz et al. 2008; Pärnänen 2016) as they are relatable to user experience. As discussed in 2.2.2 Utility there were multiple terms mentioned such as: functionality, features, and user needs concerning the factor of utility. Nielsen (2012a) explains that utility is about whether or not the system consists of the features a user needs. Therefore an interrogatory state of utility, for the respondents to answer, was created to match the thoughts of Nielsen (2012a). The attribute of performance was introduced by Lee et al. (2015b) but it concerned a system whose focus was more towards recommendations and therefore had indicators concerning errors and search quality rather than effectiveness and efficiency (Abran et al. 2003; Rubin & Chisnell 2008). Therefore, the questions were constructed in a way to better match the definition of Rubin and Chisnell (2008) and how they would measure performance. At the end of the questionnaire an open text field was also given to the respondents as a way for them to provide information about, e.g., concerns. An open text field seems to be a popular way of evaluating systems concerning user experience as it is used and recommended to apply in multiple studies (Chen and Pu 2014; Hu et al. 2015; Lee et al. 2015b).

To show consistency in music system user experience evaluation a 7-point Likert scale ranging from “very unsatisfactory/difficult poor” to “very satisfactory/easy/excellent” was used to continue in path of Hu et al. (2015) and Lee et al. (2015b). By displaying labels for the scale points it is likely to yield higher test reliability (Weng 2004). Also, Nielsen (2012c) explains that assessment of user preferences comes from a simple question on a 1-7 scale.

The background questions were not created to evaluate any type of characteristic in relation to what the respondents answered concerning the indicators in this study. The background questions were constructed to further explain that the investigation concerned the desktop client (and not other Spotify services) as well as to see if they did in fact use the desktop client. The same background questions were used in the online questionnaire as in the usability test (see 3.2.2 Background questionnaire).

For the online questionnaire the service Google Forms was used. The questionnaire described ethical considerations to better inform the respondents of the study and the usage of data collection. This was created in a similar fashion as the consent form (see Appendix 2, Consent form). The online questionnaire was created in a way to match the questionnaire of Lee et al. (2015b). Laugwitz et al. (2008) use the wording of ‘dull’ as an antonym of ‘creative’. Considering user experience studies is somewhat of a new scientific field, the researcher of this study thought it would be good with consistent uses of terms and highly regarded theoretical presence.

### 3.2 Usability tests

Throughout the years of evaluating usability, various evaluation methods have been presented; such as heuristic evaluation, cognitive walkthrough, pluralistic usability walkthrough, and formal usability inspections (Hollingsed & Novick 2007). The most effective way of understanding what
types of things that works (and what does not) in an interface is to watch people use it (Nielsen Norman Group 2014). By doing usability tests one will get close to the users and see potential errors and understand by observing and questioning (Dumas & Redish 1999). For example, performance was one of the factors displayed in 2.3 Conceptual framework and performance data is possible to collect by conducting usability tests; by measuring error rates, times, counting between elements, etc. (Rubin & Chisnell 2008). Therefore usability tests were conducted to go in depth on the various attributes (chosen for this study) but also for evaluating the performance of Spotify’s desktop client. Rubin and Chisnell (2008) explain that by testing at least four to five participants most of the usability problems will be exposed. Nielsen (2012b) states that no more than five participants are needed because this lets the researcher find as many usability problems as one would find using many more participants. By testing with a low number of participants the usability tests will aim at giving insights to drive design forward (Nielsen 2012b). On account of aforementioned statements only five usability tests were carried out.

For this study a test plan was created as it functions as a blueprint for the tests (Rubin & Chisnell 2008). Since usability tests often concerns more than just observing and task completion, the researcher thought it would be a good platform to structure the usability test in context with the study’s purpose to indicate and describe user experience positives and flaws (See Appendix 2). The test plan was outlined with points given out by Rubin and Chisnell (2008) as it mentioned:

- purpose of test and research purpose;
- participant characteristics;
- test environment and preparations, and;
- method (test design) consisting of:
  - Consent form.
  - Orientation script.
  - Background questionnaire.
  - Tasks (observation and benchmarks).
  - Semi-structured interview.

3.2.1 Pilot test
Rubin and Chisnell (2008) state that it is important to be comfortable in the testing scenario: the timings of the tests, look for design flaws in the test, read through the questionnaires, and make sure that the test is revised if needed. Schade (2015) mentions that pilot testing helps fine-tune the upcoming usability tests and thereby leading to more reliable results. Rubin and Chisnell (2008) explain that it is ideal to conduct a pilot with a real participant because that will mimic a real test and also give a good understanding of the length of doing tasks. Schade (2015) explains pilot testing is important because it is an opportunity to validate the wording of tasks, the time necessary for the session, and if it can supply an additional data point for the study. There are often multiple instruments involved in a usability test other than just being an observation of a set of tasks for a participant; and therefore a lot of things to check in a pilot test Rubin and Chisnell (2008) make the point of ambiguity being able to sneak its way in instruments such as questionnaires.

From doing the pilot test the researcher specified that the first task should not be to play a sound after finding the correct one (as it was originally planned). The reason being that screen-capturing software covered the short discussion of finishing that task. Also, the researcher specified in the orientation script – for time taking purposes – that the participant should let the test leader know when the participant has finished the task. It also came apparent that an unstructured interview was a key approach considering, in the pilot test, the participant still navigated the system while discussing according to the factors in this study.

3.2.2 Background questionnaire
A background questionnaire is created to ensure that the study’s participants are the “right” type of people (Rubin & Chisnell 2008). It is important to conduct a background questionnaire that will focus on characteristics that may influence performance (Rubin & Chisnell 2008). The background questionnaire provides historical information about the participants to better explain their behavior and performance during a test: asking questions concerning their experience, attitudes, and preferences for example.
Considering this study is concerned with Spotify’s desktop client and its users a background questionnaire was created to ask questions regarding the study participant’s experience with the service:
1. Do you have a Spotify account?
2. Have you used the Spotify desktop client?
3. How many hours per week do you spend using Spotify’s desktop client?

3.2.3 Usability task
Lee et al. (2015b) state that it needs to be considered to establish user tasks focusing on something that would be considered common such as creating playlists rather than trying to create tasks suitable for a specific dataset. Consequently, Lee et al. (2015b) explain that music systems are often evaluated to target functionality such as online radio functionality and music identification. Hu et al. (2015) state that it is reasonable to evaluate a service that generates recommendations by focusing on the uploading process for users. Therefore a logical order of fundamental tasks has been developed in this study to highlight basic functionality of Spotify’s desktop client:
1. Find the song “She Is” by the artist “The Fray”.
2. Go to the album (“How to Save a Life”) of the song.
3. Add the album to a new playlist.
4. Make the newly-created playlist a collaborative playlist.
5. Download the playlist (i.e. making it available to listen to in offline mode).

Dumas and Redish (1999) explain that a good scenario, in a usability test, is a short scenario with enough information to do the task – perhaps by using a written scenario. Therefore a scenario was created where the participant would like to add an entire album as a playlist and make it a collaborative playlist. A collaborative playlist is a playlist created to let others add, delete, and reorder tracks in Spotify.

Rubin and Chisnell (2008) explain that the performance attribute can specifically be measured by using benchmarks. An example could be expressed as “95 percent of all users will be able to load the software correctly on the first attempt”. Rubin and Chisnell (2008) explain that Successful completion criteria (SCC) is a way of measuring success by setting up benchmarks and/or timings. Considering this study is out of the business scope (i.e. competing with other similar services in terms of efficiency and effectiveness) of Spotify and more focused towards a user-centered approach there is no vision of what a well-chosen benchmark would be; in terms of the time it takes for the users to complete the tasks. By setting up benchmarks in a usability test, from a business perspective of profitability, there is opportunity for seeing improvements in terms of usability (Rubin & Chisnell 2008). However, considering the researcher of this study had no opportunity to change Spotify’s user interface, timings and completion of task was only being kept track of (in one set of usability tests) as a way to measure the performance level. In this way the researcher could see the abnormalities in time between the participants to an extent where a large period of time was being spent on the various tasks. The times were measured using a screen-capturing software.

3.2.4 Unstructured interview
Rubin and Chisnell (2008) state that subjective data, that measures a participant’s feelings or opinions of a service, is obtainable through a debriefing after the test session. Lee et al. (2015b) explain that what would benefit the research area of music systems as a whole is listening to users’ need and what they want as this will inform on how to improve the design of the systems in general. By these remarks an unstructured interview was conducted to relate to the factors presented in Figure 1; i.e. getting potential in-depth answers concerning the factors rather than just a rating scale. Bryman (2012) explain that an unstructured interview is a data collection method where the researcher has a list of topics to be covered (often called an interview guide). This solution was chosen to match the factors presented in the conceptual framework (see 2.3 Conceptual Framework); and the chosen solution was the respective interrogatory states mentioned in 3.1 Online questionnaire. This would give the participants a chance to expand on every question. The interview was recorded using a mobile phone. The interview guide was summarized into covering the following topics: learnability, affordance, feedback, aesthetics, utility, overall satisfaction and general thoughts (see Appendix 2, Unstructured
interview guide). The thought behind asking a question that concerns the service in general was that it could potentially uncover other factors of user experience in music systems as well as further details regarding the already presented factors in the interview guide.

4. Results and data analysis
The focus in this study was to describe the user experience flaws of Spotify’s desktop client usability tests (accompanied with an interview) and an online questionnaire was conducted in relation to the conceptual framework (see Figure 1). This chapter will outline the essential findings and analyze accordingly by following the structure of the conceptual framework by presenting three sections: usability, utility, and aesthetics. During the following sections the test participants (i.e. the people participating in the usability tests) will be discussed in the acronym TP (test person/participant) by an accompanied number to have the possibility to differentiate/telling them apart: TP1 (2016), TP2 (2016), TP3 (2016), TP4 (2016), and TP5 (2016).

The online questionnaire was shared with individuals who the researcher had knowledge of using a Spotify service. The people were contacted through emails, social networks, and in real life situations. Out of 27 contacted people 21 responded to the questionnaires. 20 out of the respondents answered they had used Spotify’s desktop client. All of the 20 respondents who met the requirements of the study were, according to the third background question, still using the Spotify desktop client every week.

The interview discussion has been summarized in Appendix 3 as not all information will be mentioned in the following sections; only the most essential parts that relates to the study’s purpose. It should also be mentioned that all test participants addressed that they were using Spotify’s desktop client at least two hours (or more) every week. The same background questionnaire was used in both the online questionnaire and in the usability tests to follow Rubin and Chisnell’s (2008) guidelines of ensuring that the right type of people is participating in a user study concerning systems.

4.1 Usability
4.1.1 Overall satisfaction
One of the main pillars of usability is satisfaction (International Standard Organization 1998; Dix et al. 1993; Nielsen 1994). The term overall satisfaction, as used in this study, seem to originally come from Pu et al. (2011) in researching recommender systems. In contrast to the performance attributes (i.e. efficiency and effectiveness) satisfaction is a subjective part of usability as it could differ depending on the user and their preferences.

The overall satisfaction system was an attribute that the respondents of the questionnaire felt positive about (see Figure 2). More than half of the respondents felt satisfied with the desktop client. There were only a small percentage of people who mentioned they were unsatisfied with the desktop client. In the interviews (see Appendix 3) TP2 the lack of consistency between the platforms of desktop and mobile The example mentioned, by TP2, addressed the history information functionality where desktop uses a list of songs whereas the mobile application uses a list of latest played playlist. This was also something that TP5 addressed. By handling different platforms there will be contextual constrains however it is important to preserve consistency and usability (Eisenstein et al. 2000). However, it was only TP3 who said they were unsatisfied with the desktop client as it was too much functionality not being used. One could argue that Spotify has managed to please their audience by making a desktop client that users feel satisfied with.
4.1.2 Performance

The other parts of usability, in addition to (overall) satisfaction, are effectiveness and efficiency (Dix et al. 1993; Nielsen 1994; International Standard Organization 1998). Dix et al. (1993) make an extensive explanation of effectiveness as it involves: percentage of tasks accomplished, ratio on failure of handling, and percentage of tasks achieved per unit of time. Efficiency, on the other hand, involves: time to complete the task, time spent on errors, error’s percentage, and number of available commands not called upon.

The usability tests involved time taking measures for the potential to see any abnormalities for the users to do logical tasks in the desktop client. As explained by Rubin and Chisnell (2008) efficiency and effectiveness are easily obtainable through usability testing and evaluating those attributes by measuring time it takes to complete task and the task completion itself. The five different tasks were quite easy for the participants to complete except for task 3 “Add the album to a new playlist”. It seemed like TP1 and TP5 were familiar with the existing functionality of adding an album to a playlist. However, the other three participants had difficulty in finding the correct path in the task and got easily confused as to where one should click. Although, all the tasks were completed successfully but there were definitely some of the participants who had trouble completing task 3 in terms of time (see Table 1). One could argue that this might be an issue of learnability (another attribute in this study). However, as stated by TP5 who completed the task quicker than anyone else, they had never created a playlist in that way. To conclude the usability tests, all the participants showed and understood the tasks.

<table>
<thead>
<tr>
<th>Test Person</th>
<th>Task 1</th>
<th>Task 2</th>
<th>Task 3</th>
<th>Task 4</th>
<th>Task 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP1</td>
<td>10s</td>
<td>2s</td>
<td>30s</td>
<td>20s</td>
<td>3s</td>
</tr>
<tr>
<td>TP2</td>
<td>12s</td>
<td>2s</td>
<td>2m45s</td>
<td>10s</td>
<td>5s</td>
</tr>
<tr>
<td>TP3</td>
<td>1m40s</td>
<td>2s</td>
<td>2m5s</td>
<td>15s</td>
<td>8s</td>
</tr>
<tr>
<td>TP4</td>
<td>30s</td>
<td>0s</td>
<td>2m38s</td>
<td>37s</td>
<td>14s</td>
</tr>
<tr>
<td>TP5</td>
<td>20s</td>
<td>3s</td>
<td>26s</td>
<td>14s</td>
<td>2s</td>
</tr>
</tbody>
</table>

Table 1. Participants’ time spent on the five different tasks.

The usability test showed improbabilities completing task 3 (see Table 1). TP3 believed that the system is well-designs and makes a good job of doing easy things but fails the user when it comes to more advanced tasks. Out of the 20 respondents in the online questionnaire (see Figure 3) only 5% was negative to the performance level (concerning effectiveness). The rest of the respondents felt the desktop client’s performance was positive, the answers varied between slightly good to excellent in both questions. By summarizing the three data collection methods (usability test, interview, and online questionnaire) one could argue that performance is one of the attributes that seems to be a strong part of the desktop client.
4.1.3 Learnability

One of the leading usability pioneers, Nielsen (2012a) explains that the learnability is about how easy it is for users to accomplish basic tasks the first time they encounter the design. Lee et al. (2015b) expands on Nielsen’s explanation by mentioning learnability entails ease of use and help as it should be simple, intuitive, and user-friendly. Therefore, a question in the online questionnaire was used according to Lee et al.’s (2015b) user experience study: “how easy was it to figure out how to use the system?”.

Only five percent of the respondents felt it was slightly difficult to figure out how to use the system (see Figure 4). The answers received in the interviews were positive towards the learnability of the desktop client but TP3 mentioned there was no real guidance and when using the system. By looking at the time it took for the test participants (see Table 1) to complete some of the tasks one could argue that there are a few complex processes that takes place in the desktop client that is difficult to learn. Although, it should be stated that if these participants were to use the same feature the following minute they could potentially make it quicker than before. By looking at the gathered data from the online questionnaire one could debate that Spotify users in general find the system easy to figure out. By doing multiple tests following more advanced tasks the participants might have found themselves lost in the system.

Figure 3. Bar chart illustrations generated from the online questionnaire answers concerning users’ perceptions of performance (effectiveness and efficiency) regarding Spotify’s desktop client.

4.1.4 Affordance

“How well does the system allow you to perform what you want to do?” was one of the questions designed in a prior study by Lee et al. (2015b). Affordance entails giving the user access, presenting simple functions and – in the context of music systems – being able to browse/search for specific songs/artists/albums etc. (Lee et al. 2015b).

The result gather from the online questionnaire show that the respondents have slightly more issues than the prior mentioned attributes presented (See Figure 5). TP1, TP2, TP3, and TP4 all mentioned
their concern of not knowing a lot of the functionalities or features in the desktop client. Either it was the desktop client not explaining the functions or it was not highlighted good enough to know of certain functions. This might be referred to task 3 (see Table 1) as the participants did not seem to know of the button that described more options. TP1 expressed their concern about the various sharing options and that they only stick to the sharing options that are familiar to them. TP3 mentions that the queue system seems hidden than in comparison to prior updates. TP4 explains that it is not clear enough what all the panels, elements, and icons entail.

One could argue that the system does not explain very well what the various functions are. Although, the results from the questionnaires show that the respondents felt positive about affordance attribute. Almost all of the test participants mentioned their displeasure of the system not explaining all its functions. One could argue that simpler types of tasks are much easier to understand and better explained because of their common usage. The more advanced type of functionalities seem more hidden; at according to the interviews (see Appendix 3) it seems like an indication that people would like to access them.

4.1.5 Feedback
Nielsen (1995) explains that the system should keep users informed about the current state of the system through appropriate feedback with reasonable time. Lee et al. (2015b) mention that the attribute entails clarity as the system should be clear and intuitive function and labels.

TP1 and TP5 stated that the system is communicating well. These two participants were also the ones who completed all the tasks very quickly in the usability test (see Table 3). The remaining participants (TP2, TP3, and TP4) mention that it is really difficult to understand what happens after a song, no graphical indication of what is happening, and specifically mentioning the functionality of ‘creating a playlist from an album’. According to the results gathered in the online questionnaire there was diversity in answers given from the respondents (see Figure 6). 30% of the respondents thought that the system did not do a good nor bad job of communicating. 10% of the respondents mentioned they think the system did a slightly poor job of communicating. Arguably, there are some concerns about the system not showing visibility of what the system is doing. One of the most difficult things in the usability test was creating a new playlist by adding an entire album through the album screen. By observation it became clear there was no indication of the playlist being created other than the playlist being displayed in the left panel of the system. This might be a problem depending on the amount of playlists already added, perhaps screen resolution, and interface placement. One could debate that feedback is the weakest part of the desktop client out of the collected usability indicators in this study.

Figure 5. Bar chart illustrations generated from the online questionnaire answers concerning users’ perceptions of affordance regarding Spotify’s desktop client.
Figure 6. Bar chart illustrations generated from the online questionnaire answers concerning users’ perceptions of feedback regarding Spotify’s desktop client.

4.2 Utility
4.2.1 System features matching user needs
One of the potential criterions to use in upcoming user experience studies for music systems is utility (Lee et al. 2015b). Nielsen (2012a) defines utility as something that provides the features you need. Consequently, Nielsen (2012a) states: “Does it do what the users need?”. Also, Garrett (2010) mentions that user needs is an important part of the strategic concerns in user experience. With the concept of utility being a widely used term the researcher of this study decided to create the indicator ‘system features matching user needs’ to explore the utility of the system.

One of the things mentioned in 4.1.1 Overall satisfaction was the fact of not being consistent from one platform to another. TP5 emphasized this point as it felt it would be comfortable to have the functionalities translated into the desktop client; however, it should be stated that a ‘pulse measuring’ software might not find its place on a desktop client as its purpose is made for jogging. Other participants mentioned that their needs were covered by the functionality and then some as TP2 explained that the radio functionality (i.e. displaying a radio channel based on their artist of liking) has opened up new ways for listening to music. It was definitely a positive response to the online questionnaire as presented in Figure 7. In total 80% of the respondents gave a positive response to how well the system’s features cover their needs. Potentially, as presented in the interviews (see Appendix 3), one’s only need would be to search and listen for music and that is what the responses gathers from the online questionnaires are based on. However, based on the positive responses in both the qualitative and quantitative data collection methods one could argue that Spotify does a really good job of covering the user needs. Also, as mentioned by TP2 there may be functionality that covers needs that one might not have.
Figure 7. Bar chart illustrations generated from the online questionnaire answers concerning users’ perceptions of system features matching user needs regarding Spotify’s desktop client.

4.3 Aesthetics
Lee et al. (2015b) explain that in upcoming user experience evaluation studies, of music systems, aesthetics should be one of the factors evaluated. However, the only aesthetic indicator mentioned by Lee et al. (2015b) was the indicator of attractiveness. Aesthetics is often a factor mentioned in user experience studies in words such as beautiful and/or visually pleasing but never specified to concrete indicators. Therefore this study has adopted two more indicators that have been in other multiple user experience studies (Nielsen 1995; Lavie & Tractinsky 2004; Lavie & Tracinsky 2004; Laugwitz et al. 2008; Pärnänen 2016).

4.3.1 Attractiveness
As attractiveness (or attractive) was one of the codes generated from Lee et al.’s (2015b) study it was a natural choice to use in this study – since Lee et al. (2015b) have been guiding the questions an factors used in this study. An indicator commonly referred to in user experience studies is an ‘attractive’ interface (Lavie & Tractinsky 2004; Lee et al. 2015b).

By looking at the results from the online questionnaires (see Figure 8) there definitely a diversity of answers as only ‘Very unattractive’ was the only answer not given by the respondents. In total, 60% had a positive attitude towards the desktop client’s attractiveness whereas 15% was neutral about the attractiveness and 25% thought the system was unattractive or slightly unattractive. During the interviews (see Appendix 3) TP1 suggested that a shell to Spotify’s desktop client would be a nice feature as it would enhance graphical preferences. TP3 and TP4 were very critical of the system’s attractiveness as they did not think it looked good. Arguably, there are some concerns about the attractiveness of the system as it evidently does not please all the participants of the study. This is perhaps not that surprising considering attractiveness is a subjective domain; just as music.

Figure 8. Bar chart illustrations generated from the online questionnaire answers concerning users’ perceptions of attractiveness regarding Spotify’s desktop client.

4.3.2 Clearness
Lavie and Tractinsky’s study (2004) is one of the more known aesthetic studies in as much as they have identified qualities such as: clean, clear, pleasant, symmetrical, original, sophisticated, and creative. One of the most mentionable user experience aesthetic qualities is a ‘clear’ interface (Lavie & Tractinsky 2004; Nielsen 1995). TP2, TP3, and TP4 explained that the desktop client does not have a clear structure or that it is messy. TP1 did however mention that the system did a good job of allowing them to structure playlists by creating catalogues. There was also a lot of diversity in answers given from the online questionnaires (see Figure 9). Arguably, the usability tests, and the tasks, might have provoked the participants to answer that the system being unclear because of them not completing the tasks as fast as they would like; which is a valid statement. The same participants who addressed their displeasure with the clearness of the system were the participants who had problems in completing the third task. This might indicate that clearness is correlated with the usability attribute of
feedback and could have even been potentially been interpreted as feedback to the participants. This adds to the discussion of the lack of understanding of aesthetics within human computer studies (Lavie & Tractinsky 2004). The indicator of a clear interface needs to better specified, as a ‘visually clear’ system can be misinterpreted as feedback to the participants.

Figure 9. Bar chart illustrations generated from the online questionnaire answers concerning users’ perceptions of the visual clearness regarding Spotify’s desktop client.

4.3.3 Creativity
Lastly, ‘creative’ (or sometimes innovative) design is an aspect of aesthetics that is usually mentioned in user experience or usability studies (Lavie & Tracinsky 2004; Laugwitz et al. 2008; Pärnänen 2016;). Just as the indicator of attractiveness there was a lot of diversity in the answers given by the respondents in the online questionnaire for the indicator of creative (see Figure 10). In the interviews (see Appendix 3) it became apparent that, even though the original question concerned visual creativity, the participants rather wanted to discuss the creativeness of the functions such as recommended playlists and the pulse meter adjustment in the mobile application. TP3 and TP4 stated that the desktop client looks old and dull. Also, TP1 addressed that they thought the system is not modern but that Spotify is being used by the people because of its popularity. However, creativity in a user experience music system evaluation might not be necessary to look into. For example, in the interview (see Appendix 3) almost all of the participants responded by mentioning they were satisfied with the desktop client. The creativity (visual and functionality-wise) was not as positively mentioned as overall satisfaction which might indicate that there is no real alignment between usability indicators and creativity; as seems to be the theme gathered from the interviews. It can also entail that people do not generally care for creativity in music systems; or that they care for functional creativity rather than visual creativity (e.g. displaying tracks in a specific way).
**Figure 10.** Bar chart illustrations generated from the online questionnaire answers concerning users’ perceptions of the visual creativity regarding Spotify’s desktop client.

4.4 Analysis of research design

The method of asking the same questions in the online questionnaire and in the usability tests (in the interview) was an interesting part of the study. It seemed as the usability tests made the questions more reflected in context with the tasks given during the usability tests rather than in general of the system. Presumably in the online questionnaire the questions were answered in a subjective way of the respondent’s user experience. Potentially, the online questionnaire made the participants more reflective of their own experiences. It should be stated that the interviews gave the participants a chance to expand on every aspect of the interface in contrast to the questionnaires. However, it was definitely clear that some of the participants had a negative attitude towards the desktop client. It could depend on their performance in the usability tests which seems likely by looking at Table 1 and the comments given by the participants, their possibility of elaborating on their displeasures, or simply they did not feel the system were living up to their standards even before the tests.

It should be stated that among the collected factors and indicators (see Figure 1) the factors were derived and devised from Hu et al. (2015) and Lee et al.’s (2015b) study. Two other aspects (clearness and creative) were added to this study because of their involvement in user experience studies outside of music systems (Nielsen 1995; Lavie & Tracinsky 2004; Laugwitz et al. 2008; Pärnänen 2016). After conducting the test it might be a possibility that users responded to clearness as it were feedback which might show an alignment of the two aspects. However, feedback is heavily recognizable as a way of showing visibility of an action (Nielsen 1995). Arguably, clearness could be easily defined to mean the same although it might be more appropriate to ask users about the structure as participants discussed element and panel placement. Because a system can be clear by showcasing good feedback but can also be clear by showcasing good structure. The aspect of creativity might not be a good indicator of a positive user experience. Just as the case of clearness the attribute of creativity could have been misinterpreted from a visual thing to a recommender thing as the participants rather wanted to discuss the recommendations in the desktop client rather than visual aspects.

5. Conclusions

This study’s purpose was to identify and describe user experience positives and flaws of Spotify’s desktop client. Based on the usability tests, accompanied interviews, and online questionnaires the main positives can be discussed in terms of user experience in Spotify’s desktop client:

1) The desktop client covers the needs of most users. Also, from the answers given it might also indicate that the desktop client contains features that cover needs that the user originally does not have (e.g. the radio features of the desktop client).

2) In terms of explaining and communicating basic functionality of the system (e.g. searching for and playing songs) it is often presented well. The results might indicate that more advanced functionality is not well communicated. However, by the results showed in the usability tests and explanations given in the interviews the basic functionalities are very well explained and communicated.

3) By looking at the results from the online questionnaire the performance level (i.e. the effectiveness and efficiency) of the system is positively perceived.

Also, the main flaws can be discussed:

1) The tests might indicate that the desktop client does not convey the users of advanced functionalities and when being used they are not communicated well. According to the collected data it might indicate that the fundamental functionalities such as searching for songs and playing them do not seem to be an issue. However, when presented with more advanced types of tasks there seem to be a hindrance as the desktop client does not explain where to find the functionality and the meaning (i.e. affordance). Also, when interacting with the system it seems like the more advanced functionalities are badly communicated when used (i.e. feedback).

2) From the gathered answers from the interviews and usability tests it came apparent that there are people who think of the desktop client as dull and/or unattractive. There are third party solutions to alternate Spotify’s desktop client; however, this study only covered the original version of the desktop client. Presumably – because none of the respondents or participants mentioned it – no one knew of
Spotify’s possibility of using third party solutions. This might indicate that the desktop client does a poor job of explaining its various functions and its availability to other solutions; as pointed out in the aforementioned point of affordance. One of the participant mentioned that it would be pleasant to have the availability of adapting a shell to make it look more attractive and to organize the elements of the client in a better way. This is connected with another comment that mentioned the system was messy and do not show consistency across multiple screens.

3) Multiple of the participants expresses that they dislike there is inconsistency between the platforms (desktop client and mobile application). One of the worth mentionable things were that history is not presented in the same ways on the desktop client and the mobile application. There are two different solutions to the different platforms as the desktop client shows history track by track whereas the mobile applications present latest played song in terms of what playlist the track was played in. The participant mentioned that both of these types have their advantages but it does not make sense to have the diversity between the services; even though it might be a smaller type of design.

5.1 Future studies
Lee et al. (2015b) mention that by evaluating music services one can see what users prefer in terms of design solutions and experiences. Also, by conducting researches within this area one can also see what type of aspects and research solutions to adapt. The research design of this study might have indicated that the participants of the usability tests were more reflective of their user experience of using the desktop client than the respondents of the online questionnaires. With that in mind it would be interesting to construct other tasks for users to do in upcoming user studies that concerns music systems.

From the information gathered from the different aspects in this study it might indicate that the attribute of clearness should be replaced by something more concrete such as structure. Also, when the participants were asked of their visual creative experience they rather wanted to discuss the recommendations of the system; which was an aspect discussed multiple times during the interviews. This might indicate that the recommendations are important to the user experience. In conclusion, it might be appropriate and beneficial to try to look for further aesthetics aspects in music systems as hedonic qualities have shown to be important to user experience.
References

Written sources


Garrett, J. J. (2010). Elements of user experience, the: user-centered design for the web and beyond. Pearson Education.


Oral sources

Usability test and interview datum
Appendix 1. Online questionnaire in Google Forms.
Appendix 2. Test plan

Purpose of tests and research purpose
The general purpose of this study is to focus on the commercial music service Spotify and its desktop client by describing its user experience flaws. By conducting usability tests the researcher can specifically focus on the parameter of performance (efficiency and effectiveness) by setting up benchmarks. However, by having a follow-up interview with the participants there is also an opportunity to discuss general thoughts concerning the software but also discuss the service’s utility and aesthetics.

Participant characteristics
Considering this study’s focus is to evaluate the Spotify’s desktop client the researcher thought it was a good idea to evaluate the client on current customers of Spotify. In that way it might create a good discussion of the various attributes presented in the study.

The (five) participants for this study should be people the researcher has a clue of using Spotify’s desktop client. To be sure of the participant’s experience with the desktop client a background questionnaire was created. Also, because this study involves a service which the researcher has no control over there will be no bias in the investigation.

Test environment and preparations
Because of needing Spotify desktop client-users in this study the test environment might be different for different. However, the goal is to conduct the tests in a calm environment; preferably at someone’s home or similar thus leaving the participants to focus on the tests at hand and not at a noisy environment.

- Prepare laptop for test.
- Make sure everything is reset from previous test: removed saved playlist.
- Working screen recording software to keep track of user movement – in this way a productive way of timing the participants can be made.
- Prepare mobile phone to record audio during the session.
- Prepare the multiple papers (background questionnaire, consent form, and task paper).

Method (Test design)
This section outlines the test design in a sequential structure of the usability tests.

Orientation script
An orientation script was created to show consistency between the multiple participants; thus enhancing the chances of the participants having the same conditions. The participants in this study will most likely be Swedish and therefore it is important that this orientation script will be as well.

“Hi! Thanks for participating in this test.”

“In this test you will complete a few tasks in Spotify’s desktop client. The purpose of this study is to evaluate the system of Spotify and not its users and customers: therefore there is no right or wrong answer. After filling out a short background questionnaire and completing the tasks I will ask a few questions that concerns the Spotify’s desktop client.

- Give the participants the consent form.
- After the participants has agreed on the terms given in the consent form hand over the background questionnaire.
- Give the participant the tasks on paper considering they are supposed to look for a specific item.

“During these tasks I cannot guide you”.

- After the tasks are done begin the interview questions.

Aesthetics: “How visually appealing is the system to look at?”
Utility: “How well does the system cover your needs? Are there any missing functionalities?”
General thoughts: “What are your general thoughts on the desktop client?”

- Thank the participant for its involvement/participation.
Consent form
Thanks for participating and accepting the invitation for this test.

This is a usability test concerning Spotify’s desktop client. You will be asked to:
- Fill a background questionnaire.
- Complete given tasks.
- Discuss the interface.

Participating in this test is completely voluntary. You are free to leave the test at any time without explanations. You can also ask not to be included in the test results any time in the future. All information you provide will be maintained confidentially. During the test your opinions and comments about the program will be recorded (screen-capturing and audio recording) and they may be publicly presented, reported in a course at Karlstad University called Scientific Theory and Methodology, publically presented, and shared with Spotify to improve their service. The data being collected in this test is going to be evaluated and stored in an anonymous way. Your name, or any other type of identity information, will never be used in any type of way outside of this test.

This test is not to evaluate you or judge you in any kind of way. The test is designed to measure the user experience in Spotify’s desktop client. There are no wrong or right answers to the questions being asked.

I have read and, understood and agreed on the information of this form.

Background questionnaire
To obtain data from Spotify customers, in particular those using the desktop client, a short background questionnaire should be handed out before starting the observation and tasks.
Do you have a Spotify account?
- Yes, I have my own account.
- Yes, I share an account with others.
- No

Have you used the Spotify desktop client?
- Yes
- No

How many hours per week do you spend using Spotify’s desktop client?
- Leave an open text box for the participant to answer approximately. Considering these questions are used to get an idea of the participants’ experience rather than evaluate the different types of users (such as beginners, experiences users, etc.) there is no real need to make the background questionnaire extensive.

Tasks (observation and benchmarks)
During the tasks observation and time will be examined. The test moderator will observe any abnormalities or any significant happenings. The participant will be given a paper with the task written down to ensure no language/spelling barriers are a part of this study:
Please do the following tasks:
1. Find the song ‘She Is’ by the artist ‘The Fray’.
2. Go to the album (‘How to Save a Life’) of the song.
3. Add the album to a new playlist.
4. Make the newly-created playlist a collaborative playlist.
5. Download the playlist (i.e. making it available to listen to in offline mode).

Unstructured interview guide
This is taken from the online questionnaire.
• **Overall satisfaction**: How would you rate your overall satisfaction with the system?
• **Performance**: How well does the system perform in terms of:
  o allowing you to complete tasks correctly?
  o allowing you to complete tasks time-efficiently?
• **Learnability**: How easy was it to figure out how to use the system?
• **Affordance**: How well does the system allow you to perform what you want to do?
• **Feedback**: How well does the system communicate what is going on?
• **Utility**: How well do the system’s features cover your needs?
• **Attractive**: How attractive is the system to look at?
• **Clear**: How visually clear is the system?
• **Creative**: How visually creative is the system?
Appendix 3. Summarized text from the unstructured interviews structured.

Usability

Overall satisfaction

- TP1: TP1 expresses that they are satisfied with the system because of the fulfilled needs.
- TP2: TP2 expressed their concern of the difference of functionality and look on the other platforms: track history presentation. TP2 states that Spotify makes a good effort of showcasing the recommended songs and artists even though it does not always go the way for the user. TP2 explain that they feel satisfied with the desktop client.
- TP3: TP3 express that they are not satisfied with the service because it is too much in the system that does not get used.
- TP4: TP4 states that they are satisfied with the system.
- TP5: TP5 expresses that they are extremely satisfied because of the desktop simplicity in the tasks that needs to be done.

Performance:

How well does the system perform in terms of allowing you to complete tasks correctly?

- TP1: TP1 says that sometimes the wanted track does not show up but that is mostly because of wrongly entered search string according to TP1.
- TP2: TP2 states that Spotify makes a good job of allowing the user to complete tasks effectively.
- TP3: TP3 believes that the system is well-designed to do simple things such as finding a song and playing it. However, when it comes to more advanced needs the system fails the user.
- TP4: TP4 states that it works effectively but it gives no confirmation of doing so.
- TP5: TP5 explains that they have no memory of any displeasures of completing tasks incorrectly.

How well does the system perform in terms of allowing you to complete tasks time-efficiently?

- TP1: TP1 expresses that the system is quick.
- TP2: TP2 expresses that it easy to find songs because of the advanced search functionality displaying songs in the moment.
- TP3: TP3 states that finding songs is fairly easy.
- TP4: TP4 states that is works very well to find things quickly.
- TP5: TP5 explains that everything in the system works very quick.

Learnability: How easy was it to figure out how to use the system?

- TP1: TP1 states that were no problems learning the system.
- TP2: TP2 states it is fairly easy learn things because of Spotify’s way of structuring items and elements.
- TP3: TP3 explains that by just trying to learn the system by themselves; there is no real guidance and it is difficult.
- TP4: TP4 states it is easy to follow by using the system regularly.
- TP5: TP5 explains it is easy to learn. TP5 addresses that the tasks (given in the usability tests) were easy to learn even though they had never done them before.

Affordance: How well does the system allow you to perform what you want to do?

- TP1: TP1 expresses that some parts of the system is difficult to understand because the functionalities or features are not explained. TP1 mentions the example of the various sharing options because of confusion.
- TP2: TP2 expresses that Spotify seems to have a lot of functionality that the user does not know of but by looking around for a while they seem to find more.

- TP3: TP3 explains that since the recent updates it is more difficult to understand what the system can do. TP3 brings up the example of that the queue system is hidden.

- TP4: TP4 explain that there are a few elements in the system that does not highlight its functionality. The icon (that was a part of the usability test) to explain ‘more options’ was not easily understood as it was clickable and if there was more functions behind that button. TP4 also explains that it is not clear what all the panels, elements, and icons entail at all; there should be accompanied text in most cases.

- TP5: TP5 states it is easy to understand the different functionality and they express what the functionalities are.

**Feedback: How well does the system communicate what is going on?**

- TP1: TP1 says in they only use simple functions and those are communicated well.
- TP2: TP2 states that the functionality behind creating a playlist was not well perceived at all because no graphic indication of it was happening. In general, TP2, explain that the communication can definitely be more developed.
- TP3: TP3 brings up an example by stating that it is difficult to know what happens after a song. It is not communicated clear enough.
- TP4: TP4 explains there is no confirmation of what they are doing; specifically when it came to the task to ‘create playlist from an album’ in the usability test.
- TP5: TP5 believes that the system is communicating very well.

**Utility**

**System features matching user needs: How well do the system’s features cover your needs?**

- TP1: TP1 is positive to the Spotify experience and the needs.
- TP2: TP2 feels like the system manages to match their needs and also in some ways open ups for more features such as a radio functionality (i.e. displaying a radio channel based on their artist of liking).
- TP3: TP3 expresses that the system does not make a good case of presenting its features but it works well for just listening which is the main point.
- TP4: TP4 explains that it covers the need because of the easy search and play functionality. TP4 mentions that they never look for albums but most commonly tracks.
- TP5: TP5 mentions that the mobile have a functionality that are listing tracks in a recent played playlist whereas the desktop client is listing tracks by history. TP5 states that it would be nice to have the functionality translates into the other platform; i.e. both type of presentations on both platforms. Also, TP5 mention it would be pleasant to be presented with lyrics to all songs.

**Aesthetics**

**Attractive: How attractive is the system to look at?**

- TP1: TP1 states they have a neutral feeling about the attractiveness of the system. TP1 would like to see a sort of shell to create a system that they feel is attractive.
- TP2: TP2 states that the system looks nice.
- TP3: TP3 thinks that the system is not very attractive.
- TP4: TP4 states that the system does not look good.
- TP5: TP5 thinks the system looks very good and feels modern.

**Clear: How visually clear is the system?**
• **TP1:** TP1 states that the system allows one to structure according to maps and catalogues (referring to the playlists).
• **TP2:** TP2 states that the structure is good but a few things could be clearer.
• **TP3:** TP3 expresses that the system is messy; the system and its functions should be more lucidly presented.
• TP3 thinks that the system is not consistent from interface to interface; some pictures are bigger than others and take up unnecessary space.
• **TP4:** TP4 states the system is not clear when it comes to structure. There is too much empty space in some places and in other places there is too much information.
• **TP5:** TP5 states it is extremely clear.

**Creative: How visually creative is the system?**

• **TP1:** TP1 says the system is not modern except for the fact that recommended playlist is innovative. TP1 states that people probably uses Spotify because of its popularity and not the creativeness in the system.
• **TP2:** TP2 explain that the creativeness lies in the other services at Spotify such as the mobile application. The reason being the application can adapt to an appropriate song depending on the pulse while being out jogging.
• **TP3:** TP3 believes that the system looks old.
• **TP4:** TP4 states the system is dull.
• **TP5:** TP5 believes the recommendation functionality is creative.

**General thoughts**

• **TP1:** -
• **TP2:** TP2 expressed their concern of the difference of functionality and look on the other platforms.
• TP2 states that Spotify makes a good effort of showcasing the recommended songs and artists even though it does not always go the way for the user.
• **TP3:** -
• **TP4:** -
• **TP5:** -