Technische Universität Berlin
Fakultät I – Geistes- und Bildungswissenschaften
Institut für Sprache und Kommunikation
Fachgebiet Audiokommunikation

This thesis is submitted for the degree of
Master of Science

The Impact of Choice Overload on Music Listening Experience

Melanie Schulz
[redacted]
21. März 2021

First Supervisor: Prof. Dr. Stefan Weinzierl

Second Supervisor: Prof. Dr. Jochen Steffens

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt gegenüber der Fakultät I der Technischen Universität Berlin, dass die vorliegende, dieser Erklärung angefügte Arbeit selbständig und nur unter Zuhilfenahme der im Literaturverzeichnis genannten Quellen und Hilfsmittel angefertigt wurde. Alle Stellen der Arbeit, die anderen Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind kenntlich gemacht. Ich reiche die Arbeit erstmals als Prüfungsleistung ein. Ich versichere, dass diese Arbeit oder wesentliche Teile dieser Arbeit nicht bereits dem Leistungserwerb in einer anderen Lehrveranstaltung zugrunde lagen.

Mit meiner Unterschrift bestätige ich, dass ich über fachliche Zitierregeln unterrichtet worden bin und verstanden habe. Die im betroffenen Fachgebiet üblichen Zitiervorschriften sind eingehalten worden. Eine Überprüfung der Arbeit auf Plagiate mithilfe elektronischer Hilfsmittel darf vorgenommen werden.

[redacted]

Abstract

In almost every product domain people are confronted with an oversupply. It seems that (online) shops, and thus assortments, are getting larger and larger. This is not just the case with physical goods, it also affects digital media like movies and music via streaming services. With music streaming services like Spotify and Deezer (just to name some of them), individuals get access to an enormous amount of songs. What impact does this immense amount of accessible songs have on the listening experience of the end user? The aim of this thesis is to investigate on exactly this question. In an online survey, participants will be presented with playlists of different sizes. Another manipulated factor will be the information given to the songs in these playlists. After choosing one song from each of these playlists, test subjects will have to give information on their choice-making process – in terms of difficulty, predicted satisfaction, enjoyment and frustration – and then proceed to listen to the song they have chosen. Afterwards, they will answer questions regarding their listening experience. This includes items on aesthetic properties of the music, subjective value, satisfaction and more. Results show that the negative consequences of choosing from large assortments (called *choice overload*), which have been found in prior research, are also present in very small assortments. The best choice-making experience is present in middle sized sets, until a point is reached where the negative effects outweigh the positive side of choosing from a large number of alternatives. This point is considered the threshold for choice overload. With more attributes presented to each choice option, participants found the choice-making process to be more satisfying, shifting the threshold for choice overload to larger assortments. The results further suggest that an increasing level of experienced unpleasantness in the choice-making process results in decreasing levels of the actual listening experience with the music, whereas less presented attributes attenuate this decrease.

Zusammenfassung

In fast jedem Produktbereich ist man mit einem Überangebot konfrontiert. Es scheint, als würden (Online-)Shops, und damit die Sortimente, immer größer werden. Dies ist nicht nur bei physischen Gütern der Fall, sondern betrifft auch digitale Medien wie Filme und Musik über Streaming-Dienste. Mit Musik-Streaming-Diensten wie Spotify und Deezer (um nur einige zu nennen) erhalten Individuen Zugang zu einer enormen Menge an Songs. Welchen Einfluss hat diese immense Menge an zugänglichen Songs auf das Hörerlebnis des Endverbrauchers? Das Ziel dieser Arbeit ist es, genau dieser Frage nachzugehen. In einer Online-Befragung werden den Teilnehmenden unterschiedlich große Playlists präsentiert. Ein weiterer manipulierter Faktor sind die Informationen zu den Songs in diesen Playlists. Nach der Auswahl eines Liedes aus jeder dieser Playlists müssen die Proband:innen Angaben zu ihrem Auswahlprozess machen - in Bezug auf Schwierigkeit, prognostizierte Zufriedenheit, Freude und Frustration - und anschließend das gewählte Lied anhören. Darauffolgend beantworten sie Fragen zu ihrem Hörerlebnis. Es sollen Angaben über die ästhetischen Eigenschaften der Musik, subjektivem Wert, Zufriedenheit und mehr gemacht werden. Die Ergebnisse zeigen, dass die negativen Folgen der Auswahl aus großen Sortimenten (*Choice Overload*), die in früheren Untersuchungen festgestellt wurden, auch bei sehr kleinen Sortimenten vorhanden sind. Die besten Erfahrungen bei der Auswahl sind in mittelgroßen Playlisten vorhanden, bis ein Punkt erreicht wird, an dem die negativen Auswirkungen die positiven Seiten der Entscheidung aus einer großen Anzahl von Alternativen überwiegen. Dieser Punkt wird als Schwellenwert für *Choice Overload* angesehen. Je mehr Attribute zu jeder Wahlmöglichkeit präsentiert werden, desto angenehmer empfinden die Teilnehmenden den Auswahlprozess, wodurch sich die Schwelle für *Choice Overload* zu größeren Sortimenten verschiebt. Die Ergebnisse deuten ferner darauf hin, dass ein zunehmender Grad an erlebter Unangenehmkeit im Auswahlprozess zu einer Abnahme des tatsächlichen Hörerlebnisses mit der Musik führt, während weniger präsentierte Attribute diese Abnahme abschwächen.

Contents

List of Figures

List of Tables

1 Introduction

Recent research in the domain of marketing and consumer psychology emphasizes the fact more and more, that large assortments can have detrimental causes to the decision maker. Contrary to conventional knowledge, which suggests that a vast number of options to choose from is beneficial, it shows that an increasing number of alternatives in an assortment lowers satisfaction and increases regret with the chosen alternative (Botti & Iyengar, 2004; Haynes, 2009; Iyengar & Lepper, 2000). Furthermore, the chance of deferring the choice from a particular assortment, as well as the probability of reversing the initial choice get higher (Chernev, 2003, 2005).These negative consequences of choosing from large assortments is called *choice overload*. It depicts a situation in which an individual's cognitive resources are exceeded by the complexity of a choice set (Simon, 1955). Object to these investigations have been mostly physical goods, such as wine and computer games (Hadar & Sood, 2014), jams and chocolates (Iyengar & Lepper, 2000) or camcorder and computer (Diehl & Poynor, 2010), just to name some of them. However, little research has been done in the domain of media streaming.

The need for an examination gets clear by looking at the numbers. The revenue of the music streaming market for 2021 is estimated at roughly 20 million € worldwide (Statista, 2020). With 345 million monthly active users, over 70 million accessible songs and over 4 billion playlists, Spotify displays the largest provider on the market (Spotify, 2021). Music streaming services do not just only give the opportunity to listen to known music, but also to discover new music. This can be done with the use of recommended playlists (such as the *Discover Weekly* playlist) or themed playlists, which have been created by other users.

As these playlists picture assortments as well, it is of interest to examine how the size and other traits of them influence the user's choice-making process in terms of choice overload. Research has shown, that not only the choice-making process is influenced by the number of options in an assortment, but also the outcome with the choice (Iyengar & Lepper, 2000). Therefore, the impact of the negative consequences (resulting from a large number of alternatives) on the music listening experience is another field of interest.

With being a million dollar market with such a huge amount of users, it is important to see how the user experience can be improved. To examine the impact of different sized playlists on the choice-making process and the effect of the choice-making process on the listening experience, an online listening experiment will be conducted. The design allows to examine the "perfect" size of a playlist when looking for new music, that is, the playlist size in which choice overload levels in the choice-making process are the lowest. Furthermore, the effect of the amount of presented information will be examined. After recording participant's experience with the choice-making process, the actual listening experience with the chosen option has to be rated. The results of the experiment should be able to indicate how playlists (containing new music) have to be sized and presented in order to improve the overall user experience.

1.1 Choice Overload

The term choice overload describes a situation, in which an individual's cognitive resources are exceeded by a choice problem (Simon, 1955; Toffler, 1970). This choice problem is displayed by an increase in the number of alternatives in an assortment (assortment size), which might lead to negative emotions or choice deferral. Despite conventional knowledge that consumers are drawn to large assortments and that large choice sets are exclusively beneficial to consumers (Arnold et al., 1983; Hotelling, 1929), there is a growing body of research showing that there is also a detrimental side to an increase in the number of options in an assortment (for an extensive review see Chernev et al., 2015). This could be due to the fact that people have a limited capacity to handle an increasing amount of information (Simon, 1955). Already in the 1980's there are first experimental results which show that participants experience information overload with 10 or more alternatives included in an assortment (Malhotra, 1982), with an assortment being defined by the number of options within a specific product domain (Levy & Weitz, 2006).

The outcomes of facing extensive assortments can be broad. First of all they can lead to consumers not even making a choice at all (Berger et al., 2007; Iyengar et al., 2004; Iyengar & Lepper, 2000), as well as a decrease in satisfaction with the chosen product (Chernev, 2003; Iyengar & Lepper, 2000). An induction of negative emotions can be a result of dealing with large assortments as well, such as regret and disappointment (Haynes, 2009; Inbar et al., 2011). These effects could be located in different product categories, for example chocolates and jams (Iyengar & Lepper, 2000), coffee and wine (Hadar & Sood, 2014; Mogilner et al., 2008), furniture, vacation and electronics (Chernev, 2003; Dhar, 1997) as well as retirement plans (Iyengar et al., 2004).

In further research on the choice overload hypothesis, other terms are used to describe this phenomenon: "overchoice effect" (Gourville & Soman, 2005) or "too-much-choice effect" (Scheibehenne et al., 2009). The author Barry Schwartz, who has devoted an entire book to this context, calls it the "tyranny of choice", debating that there is a seemingly endless occurrence of extensive assortments, crossing the threshold of feasible choice (Schwartz, 2000, 2004).

But there are other factors influencing the impact of assortment size on choice overload. Researchers were able to identify moderators which either attenuate or amplify the effect of extensive assortments on consumers. Those are for instance time pressure (Haynes, 2009), the number of attributes (Greifeneder et al., 2010) or subjective knowledge of the participants (Hadar & Sood, 2014).

In the following chapters, an overview of the pros and cons of large assortments as well as a conceptual model of choice overload will be presented. There will be a closer look on choice overload in recommender systems and an investigation on which of the factors dealing with choice overload can be transferred to the music streaming context.

1.1.1 Choice Overload and Assortment Size

In previous studies investigating on the choice overload phenomenon, the outcomes are considered as a function of assortment size. Accordingly, the size of the choice sets is in mostly all cases the main independent variable. The results show different tendencies though, as it is documented in a meta-analytic review by Scheibehenne et al. (2010). By looking at 50 experiments, the authors found that the mean effect size of choice overload across assortments is practically zero. Since the results apparently show advantages and disadvantages of large assortments, it is important to create an overview of the beneficial and negative consequences of large assortments. At this point, it should be pointed out that in this context only the consumer's perspective is addressed, not that of the retailer.

PROS OF LARGE ASSORTMENTS. Large assortments bring a lot of advantages for customers. An increase in the number of available alternatives in an assortment will raise the probability that consumers will find a product that complements their purchase goals (Baumol & Ide, 1956; Hotelling, 1929; Kahn & Lehmann, 1991). They also allow consumers to maintain a certain degree of flexibility, in case they are not sure about their upcoming desires and tastes (Kahn & Lehmann, 1991; Kahneman & Snell, 1992; Kreps, 1979; Walsh, 1995). When an assortment contains various options, it connotes the consumers that they have a freedom of choice (Broniarczyk et al., 1998; Kahn et al., 1987; Oppewal & Koelemeijer, 2005). Babin et al. (1994) and Botti & Iyengar (2004) document that larger assortments intensify the pleasure of shopping and the consumer's satisfaction with the chosen option. Buyers might also have a preference for bigger choice sets, as it evokes the feeling that the assortment on site consists of all potentially available alternatives (Greenleaf & Lehmann, 1995; Karni & Schwartz, 1977). Furthermore, large assortments prove to be useful to help customers with variety seeking behavior (Inman, 2001; Levav & Zhu, 2009; Simonson, 1990; Van Trijp et al., 1996; Van Herpen & Pieters, 2002). It is important to add that buyers are more likely to seek for variety in hedonic categories, rather than utilitarian ones (Van Trijp et al., 1996).

Furthermore, large assortments are helpful in situations of one-stop shopping (Messinger & Narasimhan, 1997) and they can also have educational benefits, as the consumers have the chance to learn about the range and attributes of the available products (Broniarczyk, 2008). Consumers also benefit from large assortments in terms of process-related advantages: the confrontation with multiple and new options offers satisfying stimulation (Berlyne, 1960) and raises perceived control (Langer & Rodin, 1976). They also give the possibility for the expression of individuality and autonomy and therefore add to a feeling of psychological well-being (Taylor & Brown, 1988).

These findings are cohesive with the first stage of a model by Kahn & Lehmann (1991), in which they display choice processes as hierarchical: large assortments draw consumers in at the first level of choice, whereas in the second stage, they mostly impede consumers to choose a final option. These two levels of a decision process are also called the *paradox of choice* (Schwartz, 2004). The second stage of the choice process model is supported by more recent research on the effect of large assortments on consumers, which report a negative impact.

CONS OF LARGE ASSORTMENTS. The first extensive results that showed detrimental consequences of large assortments, were made by Iyengar & Lepper (2000). They show that consumers are initially attracted by large assortments, with a subsequent difficulty to actually make a choice. The authors demonstrated that even though consumers were first more attracted to the large assortment (30 alternatives), the purchase probability (30 % vs. 3 %) was higher when tasting jams from the small assortment (6 alternatives). Another experiment of the same publication indicates that making a choice from a large assortment is more enjoyable but also more difficult. Subjects who made a choice from the large assortment also stated higher levels of regret and lower satisfaction with the chosen alternative in comparison to the small assortment group.

More findings demonstrate that an increase in options in an assortment leads to non-compensatory processing of given information, which leads to lower choice accuracy (Payne, 1976; Payne et al., 1993). Furthermore, it has been shown that satisfaction with the choice-making process decreases as the number of alternatives gets higher (Huffman & Kahn, 1998; Malhotra, 1982). Also, the level of satisfaction with the chosen product itself can be reduced due to large assortments. Satisfaction with option choice is defined as the relation between a product's performance and the expectations of the consumer (Oliver, 2003). Schwartz (2004) theorizes that large assortments raise consumers expectations about how an ideal product should perform. This means that if large assortments increase the expectations of potential buyers, they will be less satisfied with the same performance of a product, compared to a small assortment. This consideration is supported by findings of Diehl & Poynor (2010). The authors found that large assortments increase the likelihood of *disconfirmation of expectations*, which leads to a lower level of choice satisfaction and a higher chance of choice deferral. Broniarczyk (2008) speculates that consumers facing large assortments are more likely to select a status quo option in order to cope with the complexity of the choice set, which brings the possibility of experiencing regret. As the size of the choice set increases, attractive options get more and more similar which makes the choice as well as the justification of a choice more difficult (Fasolo et al., 2009; Sela et al., 2009; Timmermans, 1993). In general, choosing from a larger assortment is more connected to larger cognitive exertion in comparison to small assortments, as it requires consumers to evaluate more alternatives and more information (Haynes, 2009; Huffman & Kahn, 1998; Iyengar & Lepper, 2000). This especially contributes to consumers who have no articulated preferences. To them, larger assortments seem to be even more confusing because of the enormous amount of information that has to be evaluated to set up preferences and to eventually choose an alternative (Dhar, 1997; Greenleaf & Lehmann, 1995; Huffman & Kahn, 1998). Chernev & Hamilton (2009) and Oppewal & Koelemeijer (2005) found out that the marginal advantages of each added option decrease as the assortment gets larger.

Taken those findings into account, there are both pros and cons to large assortments. A question which raises at this point is: How many options constitute a large assortment? Looking at previous research that used assortment size as the main independent variable, there is no consistency in the choice set sizes. In the already presented study by Iyengar & Lepper (2000), the small assortment consisted of 6 alternatives, whereas the large assortment included 30 alternatives. Many of the following experiments from other authors followed these choice set sizes as a guideline and selected their assortment size conditions similarly. A common small assortment includes less than 10 options whereas the large assortments range from 24 to 36 alternatives (Berger et al., 2007; Chernev, 2003; Goodman & Malkoc, 2012; Hadar & Sood, 2014; Inbar et al., 2011). But there are experiments with other assortment sizes. Some researchers decided to make both choice sets smaller, with sizes from 1 or 2 vs. 5 alternatives (Chernev, 2005; Gourville & Soman, 2005) or 3 vs. 10 options (Haynes, 2009). On the other hand, there are authors who made the small as well as the large assortment size conditions bigger, ranging from 9 vs. 54 options (Chernev & Hamilton, 2009), 24 vs. 88 options (Chernev, 2006) to 60 vs. 300 options (Diehl & Poynor, 2010).

A problem with these results is, that the dependent variables to measure choice overload vary a lot. This makes them comparable just to a certain degree. Moreover, it has to be mentioned that the negative impact of large assortments could not be determined in all of the experiments. This is the case because moderating variables were taken into account in these investigations. Those moderators are for instance time pressure, subjective knowledge, or the availability of a dominant option, to name some of them. As it is stated before, the product domain also plays an important role, divided in hedonic and utilitarian products (Van Trijp et al., 1996). A further restriction with these experimental designs is, that the threshold for choice overload as a function of assortment size cannot be evaluated, as there are just two conditions for assortment size present in each of those studies. There are some papers that address the problem of just having two assortment conditions by including more choice sets (Reutskaja & Hogarth, 2009; Shah & Wolford, 2007). The results of them will be presented in the next chapter.

Overall, there are advantages and disadvantages of large assortments. The negative impact of large assortments on consumers could not be found in all of the studies dealing with choice overload, as there are more variables which have to be taken into account. Broniarczyk (2008) argues, that the sweet spot for choice overload is most likely a “calibration issue”, that depends on consumer, situation and product domain. Scheibehenne et al. (2010) add that while assortment size can be considered the basis of choice overload, there is no precise definition of what constitutes having too much choice. This means that there is still no definite understanding of when large assortments support choice and when they have a negative impact on choice. Contrary to the results of the meta-analytic review of Scheibehenne et al. (2010), that was not able to find a mean effect size of choice overload across assortments, Chernev et al. (2015) were able to identify a significant impact of assortment size on choice overload as they took moderating variables into account. These considerations signalize the need for a conceptual model of choice overload, which will be described in the following chapter.

1.1.2 Conceptual Model of Choice Overload

To understand how moderators and preconditions influence the consequences of large assortments on choice overload, it is inevitable to have a closer look at the possible outcomes of this phenomenon. With a meta-analysis of 99 observations, Chernev et al. (2015) were able to elaborate seven common measures of choice overload, which will be presented in more detail below. This analysis is also helpful to understand what moderates the effect of large assortments on consumer decision making. Four key factors could be identified, a detailed description of them will be part of this chapter as well. By involving those factors in the analysis, the impact of assortment size on choice overload is significant (Chernev et al., 2015). Taking both together, a conceptual model of choice overload can be constituted. The need for such a conceptual model is indicated by looking at prior research: independent and dependent variables are different across studies, which limits a cohesive understanding (Chernev et al., 2015). With the help of the analysis, the authors could bring the findings together into such a model, which provides a framework for understanding and working with the choice overload hypothesis.

OUTCOMES. Choice overload cannot be measured directly but is expressed through a number of different outcomes. Those indicators can either be *process-based* or *outcome-based*, measurable as a *subjective state* or as *observable behavior* (Chernev et al., 2015).

Changes of the subjective state are reflected in changing internal states, with the three main ones being *decision confidence*, *satisfaction, regret*. We know through research, that consumers who face choice overload experience lower levels of satisfaction with their decisions (Botti & Iyengar, 2004). Furthermore, people affected by choice overload are less likely to be confident with their chosen option (Haynes, 2009) and experience higher levels of regret after making a choice (Inbar et al., 2011). Not mentioned in the conceptual model – but also important – are *frustration* and *decision difficulty*. Consumers facing large rather than small assortments experience higher levels of frustration in the choice-making process and found the choice task to be more difficult (Iyengar & Lepper, 2000).

Scheibehenne et al. (2010) point out that it is problematic to use satisfaction as a single measure for choice overload. In most research the measure satisfaction refers to the level an individual experiences with the chosen option rather than with the choice-making process or the whole experience. If satisfaction is used as a single measure, it could lead to different answers as some people might look for new products or are prone to variety-seeking behavior. They therefore might be satisfied with the choice-making process itself, but not with their choice. Those different forms of satisfaction – outcome satisfaction and process satisfaction (Reutskaja & Hogarth, 2009) – must therefore be assessed separately. There are some authors who used this differentiation in their research, being able to track differences between both measures. The satisfaction with the choice-making task is often evaluated with the help of the question to which extent the participants *enjoyed* the decision task. In Iyengar & Lepper (2000) for example, participants found the choosing situation more pleasant with a bigger choice set, but were more satisfied with their chosen option from the small choice set.

As satisfaction is the most frequently used process-based dependent measure in choice overload research, there are further interesting results: Reutskaja & Hogarth (2009) for example used more than two assortment sizes by providing choice sets with either 5, 10, 15 or 20 alternatives, with the product being gift boxes. With this design they were able to demonstrate a curvilinear relationship between both outcome and process satisfaction and the number of alternatives. This curvilinear relationship has an inverted U-shape and displays a peak at 10 alternatives for both measures, with process satisfaction being significantly lower than outcome satisfaction, for each choice set. This finding shows again that there is a difference between those measures and that they should therefore be evaluated separately. They also imply that the level of satisfaction can also be low when there are too few options to choose from, suggesting that there can also be a reverse effect of choice overload.

The observable behavior is expressed through a series of outcomes as well. The most commonly used outcome-based measures for choice overload are *choice deferral*, *switching likelihood* and *assortment choice*. When people face choice overload, the probability of making a choice from a particular choice set decreases (Iyengar & Lepper, 2000), whereas the probability of reversing the initial choice increases (Chernev, 2003). Furthermore, when experiencing choice overload, individuals tend to exhibit a preference for small assortments rather than large assortments (Chernev, 2006).

With choice deferral being the most used outcome-based dependent variable, it is worth to have a closer look at how it is influenced by assortment size in more detail. Shah & Wolford (2007) were able to demonstrate a curvilinear (inverted U-shape) relationship between choice deferral and assortment size. In their method they included ten choice sets with the number of alternatives ranging from two to 20, in two steps. The results show that with an increase in assortment size the proportion of participants buying the product increases as well, until it reaches its maximum at the choice set size of 10 options, afterwards this proportion decreases. This supports the findings of Reutskaja & Hogarth (2009). Both show that an increase in assortment size not always leads to choice overload, but that there is a point which has to be exceeded. After this point consumers are more likely to experience the negative consequences of large assortments.

In contrast to the process-based measures, some of the outcome-based indicators are binary nominal variables: participants either defer or make a choice, the same goes for switching likelihood, while assortment choice can be assessed on an ordinal scale (e.g., small, middle, large…). As the subjective state is mostly measured on Likert-type scales, they make a more nuanced analysis of data possible, as they take into account more steps between two extremes and are comparable with each other (under the condition that they use the same scaling). Nevertheless Chernev et al. (2015) found out that the measures satisfaction, regret, choice deferral and switching likelihood are equally strong, and that they can therefore be used interchangeably. Additionally, it should be stated that the presented measures are not the only ones able to measure choice overload, but that they are the most used.

Prior analysis of research results dealing with the impact of assortment size on choice overload was not able to demonstrate a significant mean effect (Chernev et al., 2015; Scheibehenne et al., 2010). This is not surprising, as the studies used different dependent variables to test for choice overload. Also, the independent variable assortment size is in all but two observations binary, which means that the authors tested either with a small or large assortment, and as mentioned before, the sizes of these assortments differ a lot across studies. It follows that some researchers were able to demonstrate a choice overload effect with relatively small choice set sizes (Chernev, 2005; Gourville & Soman, 2005; Haynes, 2009), whereas other studies showed no choice overload in the large assortments (Chernev, 2006; Diehl & Poynor, 2010). By including moderators, Chernev et al. (2015) found a significant effect of assortment size on choice overload. They will be presented in more detail in the following paragraphs.

MODERATORS. All of the moderating variables (antecedents/preconditions) which were included in prior research can be categorized in four key factors (Chernev et al., 2015). There are *extrinsic* and *intrinsic factors* that can either amplify or attenuate the effect of assortment size on choice overload. The extrinsic factors characterize the choice set and do not alter across consumers. They can be subdivided into *task factors* and *context factors* (Payne et al., 1993). Task factors describe the structure and characteristics of the assortment and do not influence the value of the alternatives, whereas context factors do not affect the structural features. They consider the aspects that affect the value of each alternative and mark the overall attractiveness of the choice options and their similarity. In the following, task factors will be referred to as *decision task difficulty* and the context factors as *choice set complexity* (Chernev et al., 2015).

Intrinsic factors differ across individuals. They are related to the consumers personal knowledge and motivation and are therefore specific to every person. Intrinsic factors are constituted of *preference uncertainty* and *decision goal*. The first one depicts the degree of articulated preferences and how consumers are able to comprehend the benefits of each option

(Chernev, 2003). The latter describes to what extent a consumer targets to reduce the cognitive load when making a choice and the degree of willing to take other options into account (Chernev & Hamilton, 2009). How they influence the relationship between assortment size and choice overload is described in the following.

Decision Task Difficulty. The first moderator from this factor is *time constraints*. Researchers found out that when consumers have a limited period of time to evaluate the options, the cognitive effort increases, which then leads to a less systematic assessment: to non-compensatory processing (Bettman et al., 1998). As already mentioned in chapter 1.1.1, non-compensatory processing leads to lower choice accuracy (Payne, 1976; Payne et al., 1993), as well as to lower levels of choice satisfaction and a decrease in individuals confidence with the chosen option (Dhar & Nowlis, 1999; Haynes, 2009).

Another moderator from the factor decision task difficulty is the *number of attributes*. With more attributes describing each choice option, the choice set becomes more complex. Researchers have shown that an increase in the number of attributes in a choice set is more cognitively demanding and thus amplifies the negative consequences of large assortments (Chernev, 2003; Greifeneder et al., 2010; Hoch et al., 1999). Keller & Staelin (1987) demonstrated an inverted U-shape for the relationship between the number of attributes and choice accuracy, with an increase in choice accuracy with a rising number of attributes until the maximum is reached and the choice accuracy drops after a certain number of attributes is exceeded.

The *presentation format* is a further moderator of choice overload. When options are presented in an organized order, it decreases the difficulty and increases the satisfaction levels of consumers when choosing from large assortments (Diehl, 2005; Diehl et al., 2003; Hoch et al., 1999; Mogilner et al., 2008). If the options are presented by attributes (categories) rather than by alternatives, it makes it easier for individuals to compare and thus increases satisfaction in the choice-making process by lowering the cognitive effort (Huffman & Kahn, 1998; Mogilner et al., 2008).

Choice Set Complexity. The existence of a *dominant option* (a superior option) has an influence on the impact of assortment size on choice overload, such that the likelihood of purchasing from an assortment is higher when it includes a dominant option and thus attenuates choice overload in large assortments (Boatwright & Nunes, 2001; Broniarczyk et al., 1998; Chernev, 2006; Oppewal & Koelemeijer, 2005). If all options are more similar to each other, the chances of deferring the choice from a greater assortments are higher (Dhar, 1997; Tversky & Shafir, 1992).

The *attractiveness of choice options* is another moderating factor belonging to choice set complexity. Attractiveness is defined by the quality of the choice options. Chernev & Hamilton (2009) were able to show a preference for smaller assortments when they consisted of more attractive options. An explanation could be that if all options are attractive, they are more similar to each other and make a comparison more cognitively tasking and therefore weaken the preference for large assortments.

A further moderating effect is the *alignability of given attributes*. Options with alignable attributes are defined by including one feature on different levels, whereas nonalignable attributes are not easily comparable as they consist of different features (Markman & Medin, 1995). When the alignability of attributes in a choice set is high, the probability of purchasing from this choice set increases, and hence attenuates choice overload (Gourville & Soman, 2005).

Preference Uncertainty. The *expertise* of an individual affects the negative consequences of large assortments as well. It has been shown that a low level of expertise is connected to a greater likelihood of deferring the choice from large assortments, while individuals with high expertise in the given product domain tend to exhibit a preference for large assortments and are more likely to purchase from them (Chernev, 2003; Mogilner et al., 2008; Morrin et al., 2012).

Contrary to the real expertise, *subjective knowledge* (consumers' knowledge compared to an expert or amateur group) shows a reverse effect: consumers who feel knowledgeable in the product domain (high subjective knowledge) would rather purchase from a small assortment and vice versa (Hadar & Sood, 2014).

Decision Goal. This factor is constituted by *decision intent* and *decision focus*. Decision intent describes the degree to which consumers aim to browse through an assortment or whether they actually want to make a choice. Choice overload is reduced when individuals rather have a browsing goal than a buying goal, as it decreases the cognitive costs and the responsibility (Chernev & Hamilton, 2009).

Decision focus describes whether consumers intend to choose among assortments or an item from an assortment at hand. Individuals who choose among assortments show a preference for large assortments, whereas choosing a product from a given assortment increases the cognitive costs and makes the choice more difficult, which results in a preference for smaller assortments (Chernev, 2006; Huffman & Kahn, 1998).

Additional to the intrinsic factors and decision goal, there is another moderator that is specific to each individual. That is the *propensity to maximize*. According to Schwartz et al. (2002), people can be categorized into two groups: maximizers and satisficers. Maximizers are individuals who will always strive for the best option within a given choice set, whereas satisficers are content with an option that exceeds the threshold of acceptability, something that is good enough (Simon, 1955). It has been demonstrated that maximization correlates negatively with happiness, optimism and satisfaction, and positively with perfectionism and regret (Schwartz et al., 2002). The authors also found out that in consumer decisions, maximizers experience a lower level of satisfaction than satisficers. The findings lead to the assumption that the propensity to maximize or satisfice of individuals effects the impact of assortment size, as it is strongly linked to choice behavior and the possible outcomes of choice overload.

With their meta-analysis, Chernev et al. (2015) were able to show that all of the moderating factors are statistically significant and that they have a powerful effect on choice overload. Furthermore, they show that all of the moderators produce similar effects on the outcome measures. With the help of the analysis, a good conceptual model for choice overload could be created, which perfectly brings together the possible outcomes and moderators. It is able to explain the construct in detail and illuminates what actually constitutes “too much choice”. An overview of the key factors and the measures is presented in Figure 1.

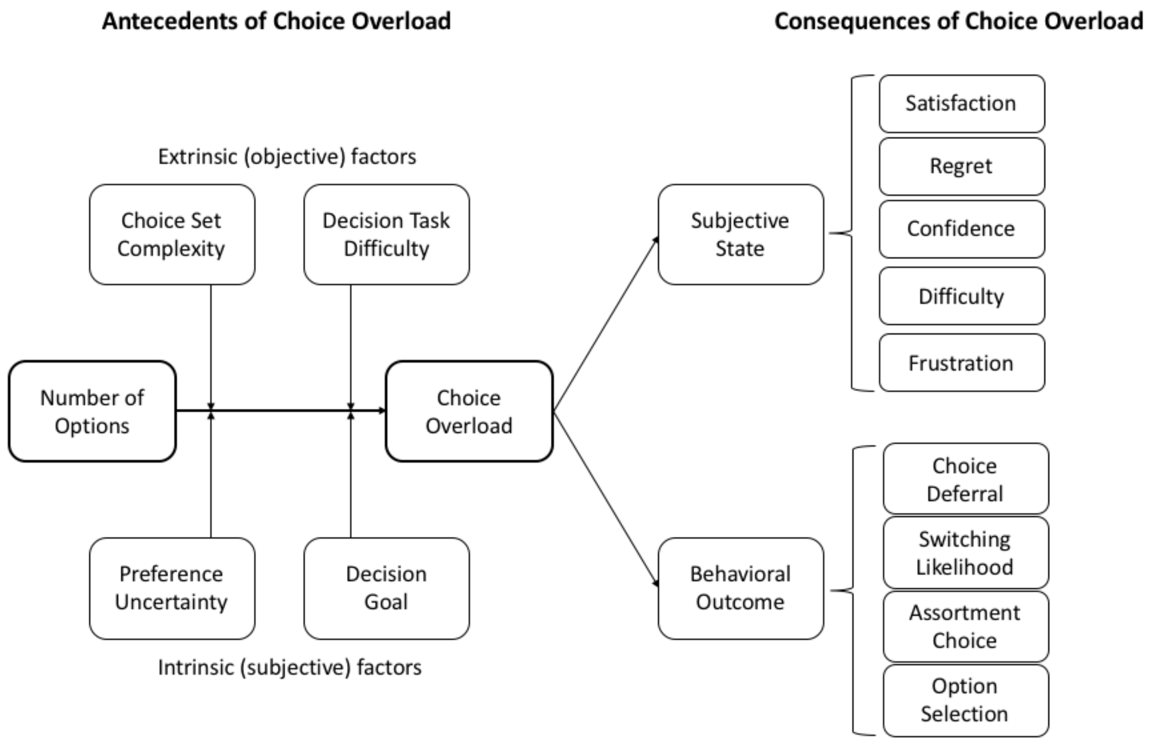


Figure 1. *Conceptual model of choice overload as a function of assortment size (adapted from Chernev et al. (2015)).*

In the following chapter, two studies will be presented which deal with choice overload in recommender systems and include music as the product domain. What follows is an investigation on moderators of choice overload, and how they can be applied to the music streaming context.

1.1.3 Choice Overload in Music Streaming

Music streaming services often give the possibility to their consumers to browse through new music through the creation of recommended playlists, such as the *Discover Weekly* playlist by Spotify. Based on prior music choices and plays, an algorithm creates an individual playlist for each user containing 30 recommended songs. Additionally, users have the chance to look for new music with the help of themed playlists created by the provider or other users. As such playlists display assortments with a certain number of alternatives to choose from, the probability of experiencing choice overload, hence influencing the consumer's subjective state and behavioral outcome, is existent. Given this knowledge, it would be interesting to see how the number of options in a playlist affects the consumers' choice-making behavior and how it influences the experience with the chosen option (the *listening experience*). To better understand the impact of assortment size on decision making in the music streaming context, two studies are going to be presented and discussed in the following paragraphs. At the end of this chapter, the moderators of choice overload will be evaluated on how they can be applied to and implemented in this context.

CHOICE OVERLOAD IN MUSIC. As mentioned before, a lot of several product categories have been the object to choice overload research, and Broniarczyk (2008) argues that the threshold for choice overload is influenced by the product domain. Therefore, the product category can be seen as a moderator as well, as the type – whether utilitarian or hedonic – has an impact on the negative consequences of large assortments as well (Babin et al., 1994; Sela et al., 2009; Van Trijp et al., 1996). Van Trijp et al. (1996) predict a preference for larger assortments in the hedonic product domain, as people tend to experience a higher level of variety-seeking behavior with this product type. Contrary to this, Sela et al. (2009) report that consumers rather choose utilitarian products from large assortments, as they are easier to justify. Even though music can have psychological functions (for an overview see Schäfer et al. (2013)), it is more suitable to categorize it as a hedonic product as it provides emotional gratification, experiential consumption, pleasure and excitement (Hirschman & Holbrook, 1982). Although music is a good being consumed by a wide population with millions of songs being widely accessible, it is rarely dealt with in research on choice overload. Scheibehenne et al. (2009) chose classical music recordings as the product in one of their studies. To see whether choice overload can be detected in the music context, the methodology and the results will be discussed.

The authors used several moderator variables: propensity to maximize, consideration set size (the number of options considered before making a final choice), search behavior, expertise, option attractiveness and cultural differences. The within-subject design consisted of two assortment sizes of either 6 or 30 options, containing classical vocal or orchestral music CDs published by the label *Deutsche Grammophon*. Each of the participants chose from two sets with four combination possibilities from *assortment size* x *musical style*. The information given to each CD were 4 attributes: the picture of the cover, the last name of the composer, the orchestra/choir and the conductor. Without time constraints, participants were able to look at all the options, and after choosing one of them they listened to a two-minute sample. The recorded dependent measures (measures to record choice overload) included post-choice satisfaction, post-choice regret and choice motivation.

A main effect of assortment size on the dependent variables could not be found, as around 50 % of the participants were more or equally satisfied with choosing from the large set rather than the small set, the same goes for the variable post-choice regret. The choice motivation was a little higher (64 %) in the large assortment size condition. A difference concerning the musical styles was not present either. Regarding all the moderating variables, no effects were found as well.

Scheibehenne et al. (2009) argue that more than one manipulation (not just assortment size) is necessary to extract a main effect of assortment size on the different outcomes, which would decrease the generalizability of a possible effect though. The results confirm the findings of Chernev et al. (2015), that preconditions need to be present to obtain a significant main effect of assortment size on choice overload, although the precondition *expertise* had no influence on the outcome. An explanation could be that the expertise of the sample did not vary enough to possibly influence the outcome, as it has not been controlled and evaluated before the experiment. However, no information is given about the range of expertise of the participants, as the distribution is not presented in the paper. Apart from the lack of other preconditions, one could assume that a choice overload effect might be absent in music, or that the choice sets need to be significantly larger to be able to detect am impact. Looking at the amount of accessible songs nowadays, it could be that individuals became accustomed to large choice sets in the music domain and that choice set sizes need to be adapted in future experiments. Furthermore, the choice of a CD or a song requires less responsibility rather than choosing an MP3-Player, microwave or other utilitarian products as it is quite ephemeral, and – when using a music streaming service – there are no additional costs. Also critical in this context is that the experience of the choice-making process was not evaluated, as just post-choice satisfaction and regret were assessed. Maybe some participants experienced the choice to be frustrating and not enjoyable but were satisfied with the chosen option after listening. This assumption is a further reason to separate those stages from each other.

CHOICE OVERLOAD IN RECOMMENDER SYSTEMS. Additional to themed playlists, a tool to search for new music can also be recommended playlists which have been created algorithmically. As such playlists contain a certain amount of songs as well, it is important to have a look at how the size of such playlists affects the consumer. As these playlists are based on the individual preferences, the attractiveness of the options plays an important role in this context and can possibly be a crucial moderator. As mentioned in chapter 1.1.2, a small assortment containing attractive options is preferred to a large one, as because of the attractiveness, the similarity increases and makes trade-offs more difficult (Chernev & Hamilton, 2009). Bollen et al. (2010) executed such an experiment with movies as the product, with assortment size and option attractiveness being manipulated. The method and results will be presented and discussed in the following.

First, the authors had to record participant's preferences by letting them rate ten movies they are familiar with, in order to offer options of high quality to each of them. Afterwards, every participant was presented one of three movie sets. First one containing five options, all of them recommended by the algorithm (high quality); second one containing 20 options with movies also recommended by the algorithm or the last one with 20 options including the Top 5 recommendations and 15 other movies not matching the preferences. The recorded dependent measures were perceived variety, set attractiveness, choice difficulty and satisfaction with the chosen item.

The results show no differences in satisfaction levels between the choice sets. The smallest set containing five options had the lowest ratings on perceived variety, as well as on difficulty, whereas the large high-quality set offered the most variety but was the hardest to choose from. The large, low quality set was rated high on perceived variety and low on difficulty and attractiveness. The results support previous findings, that when adding attractive options to a set, the choice difficulty increases (Chernev & Hamilton, 2009). The authors argue that the inverted U-shape of satisfaction demonstrated by Reutskaja & Hogarth (2009) might be applied and that a set size between five and 20 options brings the best compromise of variety, attractiveness and difficulty. This reflection emphasizes the notion that more set sizes should be looked at in future experiments. A limitation with the satisfaction measure – and hence no significant differences between the sets – could be that participants were not able to watch the full movie, making it harder to assess satisfaction with the chosen option. And besides the measure of choice difficulty, participants could not rate their actual choice-making experience, making it hard to evaluate whether choice overload occurred or not. This point reinforces the idea of considering the choice-making process and the outcome of a choice separately.

None of the presented studies was able to show a choice overload effect in each of the conditions. Even though previous findings regarding the influence of option attractiveness on choice difficulty could be supported, the satisfaction ratings were similar across assortments. The causes for not being able to demonstrate a choice overload effect can be several: First, both studies implemented the independent measure assortment size as a binary variable, with a small and a large assortment, and a big gap in the middle. If process and outcome satisfaction are really a function of assortment size in form of an inverted U-shape, the satisfaction peak could lie in between the chosen assortment sizes. Another consideration is that the peak of satisfaction could also be at an assortment size above the chosen experimental conditions, as music and movie streaming services already offer an enormous amount of options to choose from, hence consumers got used to a vast number of alternatives. As Scheibehenne et al. (2009) point out, other preconditions might have to be existent as well. Looking at the dependent measures of these studies, the authors evaluated post-choice satisfaction (satisfaction with the chosen option) and post-choice regret, but no real assessment of the choice-making process (besides choice difficulty) took place. This might be a problem, as the choice process itself could be satisfying to individuals, whereas the chosen option might be unsatisfying.

The results suggest that future experiments, dealing with choice overload in music streaming services, have to take more assortment sizes into account, as well as larger assortments. Further preconditions which are suitable to this context should be considered too. Moreover, it seems important to separate the evaluation of the choice-making process from the outcome with the choice (the experience of the users with the chosen option – from this point on referred to as *music listening experience*), making it also possible to see whether these have an interaction. To decide which other preconditions can be used to develop a suitable experimental design to test for choice overload in music streaming services, the conceptual model is going to be analyzed in the following.

PRECONDITIONS OF CHOICE OVERLOAD IN MUSIC STREAMING. To create a suitable experimental design to see whether and how choice overload can be existent in the music streaming context, a further look on the moderators of the presented conceptual model is substantial. What follows, is the presentation of the moderators and a derivation of the relevance.

Time constraints are not relevant in this context as, in a realistic scenario, individuals probably do not feel rushed when selecting music from playlists. A manipulation is therefore not needed.

The *number of attributes* is of interest, as music can be presented with different amounts of information (as in Scheibehenne et al. (2009)). In both the mobile and desktop version of Spotify for example, just the title and artist of the songs are presented in a playlist. As music can be described by more than two attributes, e.g. the musical style or adjectives describing the music, it would be interesting to see how a change in number of attributes affects the impact of assortment size on the consumer. Jacoby et al. (1974) theorize that the amount of information in an assortment to be processed is the product of the number of alternatives multiplied by the number of attributes. Following this idea, an increase in the number of attributes could lead to more cognitive effort and can thus amplify choice overload. Greifeneder et al. (2010) were able to support this theory, as the choice sets with a higher number of attributes (6 attributes) were perceived as more complex compared to the choice sets in which the options were just described on one attribute. The rating of choice satisfaction was non-significant in the analysis, once again, supporting the notion that the choice-making process should be evaluated separately from the choice outcome. To see how the number of attributes influences the negative consequences of large assortments in the music streaming context, a manipulation should be included in future experiments.

As playlists must be visually presented in some way to make a choice possible, the *presentation format* might be relevant as well. It has been shown that an organized order of choice options attenuates choice overload (Diehl, 2005). In this context, an "organized order" could be categories (genre, album, artist, etc.). Mogilner et al. (2008) report an increase in satisfaction levels of choosers who are unfamiliar with the product domain, when categories are present. As categories are usually not present in playlists (as the playlist itself depicts a category), this precondition does not need to be manipulated. Instead, it should be held constant by a randomization of the alternatives for each participant.

The existence of a *dominant option* increases the likelihood of purchasing from a large assortment and therefore decreases choice deferral (Chernev, 2006; Oppewal & Koelemeijer, 2005). Such a dominant option can be songs that participants already know or like, like popular songs. To manipulate this factor, observations of participants preferences and knowledge of songs would have to be done before an experiment, which would be hard to implement. Instead, it should be controlled by choosing unfamiliar stimuli to rule out the possibility of the participants knowing the songs. As there is still a chance of familiarity with the options, a question regarding the knowing of a chosen alternative should be included.

As Bollen et al. (2010) were able to show, the *attractiveness of the choice options* has an impact on choice difficulty. As these results support previous findings, there are other more interesting preconditions which have not been tested in this context yet. Additionally, to offer attractive options to participants in an experiment, the preferences need to be known. No manipulation is needed, but it should be controlled by asking participants how they would rate the perceived quality of the choice options.

The *alignability of given attributes* might play a role in the music streaming context. Non-alignable attributes are not realistic in a playlist scenario though, as playlists often are often presented in a tabular form, so that all songs are described on the same attributes but having different levels of these attributes, making them alignable. Therefore it does not need to be manipulated.

The *expertise* of participants was already suggested to have an impact in the music context by Scheibehenne et al. (2009). However, the authors were not able to detect an effect. To gain significant data, the participants would have to be tested in expertise levels before an experiment, in order to divide them in a control and an experimental group. Therefore a manipulation of the variable can be omitted, as it is too complex to implement and to find suitable participants for it.

Both moderators *decision focus* and *decision intent* might be relevant as the possibility to choose among assortments (playlists) is highly realistic. But in line with the presented studies, the focus should be more on the negative consequences of choosing a single option from large assortments, rather than choosing among assortments.

As the above discussion suggests, the *number of attributes* proves to be an interesting precondition, which is easy to manipulate. Other moderators such as *dominant option*, *attractiveness of options* or *presentation format* are likely to play a role in the music streaming context as well, but they have either been dealt with in previous research or are harder to manipulate. Therefore, the *number of attributes* marks a good start for investigating more on the choice overload phenomenon in music streaming. The *propensity to maximize* should be recorded in any case, as it is a measure that is individual-specific and likely to influence the outcome.

At this point, it should be stated that the focus of interest are playlists, which have not been organically created by the user itself, instead by other users or an algorithm. A hypothetical situation for example is that users want to listen to new music. How music choices influence the listener is going to be presented in the following chapter.

1.2 The Impact of Music Choice

Additional to the choice overload hypothesis, which illuminates the impact of assortment size on individuals, it is necessary to have a look at how music choices influence listeners. Even though research on choice overload covers subjective outcomes and preconditions, it is important to understand how music choices influence the listening experience from a music psychological perspective.

In research regarding music choice and its consequences, a distinction is made between individuals not having a choice over the music selection or them choosing the music. It has been shown that the emotional response to music is more pronounced when it is chosen by the listener itself instead of by others or if it has been randomly selected (Blood & Zatorre, 2001; Liljeström et al., 2013; Mitchell et al., 2008). When individuals reported higher levels of choice, an increase in enjoyment, concentration functions and relaxation could be observed, whereas no such effects could be demonstrated with participants reporting low levels of choice (Greasley & Lamont, 2011). Moreover, higher levels of choice support positive mood change, suggesting that self-chosen music mitigates negative emotions (Sloboda, 2010).

The level of choice is in this context equated with the level of control. Krause et al. (2014) were able to show when individuals experience a low choice level, they perceive the situation as not having control over the music choice, resulting in negative effects. Furthermore, they have shown that having control of the music selection results in positive outcomes like motivation and enjoyment. An increase in intensity of positive emotional reactions is also reached by selecting music one prefers (Liljeström et al., 2013).

Even though there are no specific findings on the impact of the number of available options (songs, playlists, etc.) on listening experience, the presented findings could be helpful in creating a connection between music choice and choice overload. When confronted with playlists to choose an item from, the consumer experiences a high level of choice – she/he has to make the choice her-/himself. The number of options could influence the perceived level of control though. And as choice overload is linked to low levels of satisfaction and enjoyment (Iyengar & Lepper, 2000), and low levels of enjoyment to not having control about music choices (Krause et al., 2014), one might suggest that an increase in the number of options in a choice set could decrease the perceived level of choice/control, hence induce negative emotions. This could lead to a decrease in the overall listening experience.

On the other hand, an increase in the number of attributes (e.g., the adding of the genre or musical style) could raise the level of perceived control, as listeners would be able to choose something they prefer, which also contributes to positive emotions (Liljeström et al., 2013). When looking for new music and just the title and artist are presented, consumers have no clue what to expect.

In the following chapter, suggestions on how all of the presented constructs and factors could influence choice overload and the listening experience will be stated, as well as the relevance of an investigation on these suggestions.

1.3 Hypotheses and Relevance

This study focuses on the negative consequences of large assortments on choice overload and the listening experience. Looking at the presented literature and its limitations, the discussed points are most important:

- More than one manipulation is necessary.
- A threshold for choice overload could not be found, as most studies examine just two choice set sizes.
- There is a need for larger assortments to test for choice overload in the music domain.
- A separate evaluation of the choice-making process and the outcome of the choice (the listening experience) is necessary, also to see whether the first has an impact on the latter.
- The attractiveness of choice options and the existence of a dominant option moderate the impact of assortment size on choice overload, and thus have to be controlled.

These aspects should be addressed in the experimental design. By including this knowledge, it should be possible to see:

a) how the size of an assortment and the number of attributes influence the choice-making process of an individual.
b) at which assortment size the negative consequences of large assortments are the weakest/the strongest.

c) if choice overload (in the music domain) occurs in larger choice set sizes than examined before.
d) how the consequences of the choice-making process influence the listening experience.

Taken the findings of prior research into account, the following results are expected:

1. An increase in assortment size and number of attributes will result in higher levels of choice overload (difficulty, frustration and switching likelihood). However, there will be a peak of enjoyment and anticipated satisfaction in the choice-making process at a certain set size, which will not be the smallest and the largest.
2. An increase in choice overload leads to negative consequences in the outcome with the choice, hence the listening experience.

2.a) With a high number of attributes, the music listening experience will generally be higher compared to the listening experience participants face with a low number of attributes. A negative linear relationship will be present for both conditions, however, the slope of this relationship will be less steep in the high informativeness condition.

3. Higher levels of propensity to maximize amplify choice overload and thus mitigate the listening experience.

Because the influence of the attractiveness of options and the existence of a dominant option should be checked as well, further results which are not part of the main experiment are expected:

4. If the perceived attractiveness of the options across assortment sizes differs, than assortments which have a higher perceived attractiveness amplify the choice overload effect, but also the listening experience.
5. The existence of a dominant option attenuates levels of choice overload and increases the listening experience.

The purpose of the current study is to understand when large assortments might trigger negative consequences in the choice-making process and the experience with the product. With the digitalization of the music industry and an ever-growing amount of accessible songs online via music streaming services, it is important to investigate on the user experience and to find possible limitations in order to be able to improve it.

2 Method

To gain empirical data to determine if there is a connection between assortment size and number of attributes on choice overload and music listening experiences, an online listening experiment with an overall of 75 songs was conducted. The execution of the experiment as an online survey proved very useful during the COVID-19 situation as the time of the opening of university and laboratory premises was not certain, and it mimics a realistic scenario as people do it in an environment in which they normally listen to music.

In the following sections, the participants, the stimuli, the design and measures, and the implementation and execution of the experiment will be described in more detail. In the final section of this chapter an analysis of the condition assignment and the examination of the English level will be presented.

The stimuli, experimental design and measures were selected in agreement with a fellow student, Miguel Angel Reyes Botello, whose master thesis examines the impact of assortment size on choice overload in the music domain as well but with different moderators and a slightly different experimental design. The similarity of the designs is intended to compare and possibly merge the results.

2.1 Participants

The acquisition of the participants was mainly done via E-Mail through the university, internet forums, social media and friends, family and colleagues by sending the link to the online survey. Participation was voluntary and unpaid.

In total, 170 people started the survey, of which 101 completed it. The youngest participant was 21 years old at the time of the survey, whereas the oldest had an age of 65 (mean age = 31,7, SD = 9,8). Regarding the gender of the sample, 47,5 % of the participants have indicated that they are female, 50,5 % male, and 2 % did not want to specify their gender.

2.2 Stimuli

The stimuli for this experiment were taken from an existing pool of 549 songs, which have been selected by Lepa et al. (2020) for an online listening experiment.

In their work, Lepa et al. (2020) address the problem that a wide range of music titles is available for audio branding purposes from different archives, and that the information coming with them is often inconsistent. The authors thus aimed to develop a computational model to predict perceived musical expression as a tool for audio branding experts to gain a suitable preselection for branding scenarios. For their online listening experiment, they selected 549 songs from an archive of over 100.000 titles which have been made available by a collaborating audio branding agency. This selection consisted of tracks from popular music, dance music and classical music and was categorized into ten different genre and 61 musical styles. Those 549 songs contained nine songs of each musical style, all of which have been extensively discussed by audio branding experts (Lepa et al., 2020). 30-second excerpts of these tracks were presented to 10.144 participants from Germany, the UK and Spain and had to be rated by means of the GMBI (*General Music Branding Inventory*)[1] on a 6-point-scale. The degree of familiarity had to be rated by the participants likewise on a 6-point-scale.

The preselection of songs by Lepa et al. (2020) comes with the title and artist as well as GMBI and familiarity ratings for each song. They have also already been categorized into musical styles. They therefore are ideal for this work, as the experiment includes a *high informativeness* condition (see section 2.3.1), in which the musical style and the three GMBI adjectives with the highest ratings of each song will be part of the presented playlists. 75 tracks were selected from those 549 songs. The selection procedure was as follows:

As prior research has shown, people tend to select their most preferred option (a dominant option) from a choice set that exceeds their cognitive load (Boatwright & Nunes, 2001; Broniarczyk et al., 1998; Chernev, 2006; Oppewal & Koelemeijer, 2005), assuming that if songs are known to participants, they would choose them. Following this, all songs with familiarity ratings of ≥ 2,00 (the song with the highest rating has a value of 5,44 whereas the song with the lowest rating is at 1,15) were eliminated, as the songs should be as unknown as possible to rule out any bias of the participants (e.g., that they choose a song they already know because they like it). Furthermore, the scenario of a recommended or themed playlist should be imitated, where people usually try to find new songs. The eliminating of those songs left a total of 358 tracks for further selection. With the first step, five musical styles were automatically excluded as the songs representing those had comparatively high familiarity ratings. The styles which have been excluded in this proceeding were *Mainstream*, *AOR*, *Classic-Rock*, *Rock & Roll* and *Schlager*. The aim was to have 50 musical styles left because in the *large assortment size* condition there is a maximum of 50 songs, and these should be set up as diverse as possible (each song – a different musical style). The diversity of the stimuli is important because the

[1] The GMBI was initially developed by Steffens et al. (2017) as an instrument which is able to describe dimensions of musical expression in branding contexts with the help of attribute scores.

distinctiveness between options in a choice set is connected to the perceived variety of it. It is smaller when an assortment contains more similar options (Hoch et al., 1999; Van Herpen & Pieters, 2002). It is therefore necessary to keep the option similarity/distinctiveness in the assortments constant by just using one song from each musical style, to make sure that it does not influence the outcome. Six more styles had to be ruled out as 56 were still left. To achieve this, the songs of the individual musical styles were grouped, and a familiarity rating was determined for each style by calculating the mean value from the individual familiarity ratings of the songs. The three musical styles with the highest familiarity rating (*Historical-Classical*, M = 1,98; *Classical-Jazz*, M = 1,92; *Smooth-Jazz*, M = 1,88) and the three musical styles with the lowest familiarity rating (*Oriental*, M = 1,44; *Afro*, M = 1,47; *Folkloric*, M = 1,54) were excluded from further proceedings. With 50 musical styles left, the song of each style with the lowest familiarity rating was selected for the *large assortment size* condition.

For the *small assortment size* condition there is a maximum of 25 songs. For this case, the song with the second-lowest familiarity rating was selected from every second musical style, with the styles arranged in alphabetical order. The selection of the songs with their familiarity rating, musical style, the three highest-rated adjectives (the adjectives *male* and *female* were not considered, as it in most cases reflects the gender of the singer) and whether they are part of the *small* or *large assortment size* condition is presented in Table 1.

The selected tracks had to be edited for the online listening experiment. Following the experimental design of Scheibehenne et al. (2009), investigating on choice overload with music as the product, the tracks were shortened to two-minute excerpts, starting from the beginning of the piece. The length of the stimuli should be ideal with two minutes, as it is long enough to let participants get into the music, but also short enough that the participants do not become inattentive.

Subsequently, the excerpts were normalized in loudness. As in this experiment a simulation of a music streaming platform is desired, the regulations of Spotify were adapted. The provider states on its FAQs page for artists, that they apply either negative or positive gain on the uploaded masters so that the loudness level will be roughly at -14 $LUFS_i$ [2] (Spotify S.A., 2020). To measure the original loudness of the excerpts, a VST-Plug-In (*Youlean Loudness Meter*[3]) in combination with the digital audio workstation *Ableton Live 9*[4] was used. After the measurements, a limiter (*Limiter N°6*[5]) was applied to the excerpts with the settings adjusted

[2] integrated LUFS; indicates the measured loudness over the whole song/excerpt
[3] https://youlean.co/youlean-loudness-meter/
[4] https://www.ableton.com/
[5] https://www.tokyodawn net/vladg-limiter-n6/

that the loudness resulted in around -14 $LUFS_i$ and the true peaks were at a maximum of -1 dB. To ensure whether these values were actually achieved, the excerpts were measured again after editing. Finally a fade-out was added to each excerpt and they were MP3-encoded (Stereo, 170-210 Kbit/s).

Table 1. *The complete list of stimuli with the information from Lepa et al. (2020) and their assortment size condition assignment.*

Title	Artist	Musical Style	Familiarity Rating		GMBI		S/L
Amsterdam – Paris	Lilium	Ambient	1,44	chilled	futuristic	creative	L
Koun K'teuy	Pan Ron	Asian	1,15	unique	authentic	happy	L
Balearic Incarnation	Dølle Jølle	Balearic	1,32	relaxing	chilled	modern	L
Çokollata	Fanfara Tirana	Balkan	1,53	happy	authentic	playful	L
White Picket Fence	Jesper Munk	Blues	1,78	authentic	modern	urban	L
Riot	Joe Bataan	Boogaloo	1,63	happy	friendly	retro	L
You Can't Turn Me Away	Sylvia Striplin	Boogie	1,68	chilled	soft	relaxing	L
My Heart Said (The Bossa Nova)	Irene Reid	Bossa-Nova	1,47	retro	happy	friendly	L
1, 2, Go	Icicle	Broken-Beats	1,63	modern	futuristic	urban	L
Cut Your Coat According To Your Size	Apolos Empire Rhythm Orchestra	Calypso	1,84	happy	friendly	authentic	L
Kangourou	Charles Trenet	Chanson	1,61	friendly	classic	retro	L
Wuppertal	Grandbrothers	Contemporary -Classical	1,65	relaxing	chilled	soft	L
Just As Scared	James Yorkston	Contemporary -Folk	1,50	soft	relaxing	chilled	L
That's What's Up	Edward Sharpe & The Magnetic Zeros	Country	1,64	happy	retro	friendly	L
Bam Bam	Sister Nancy	Dancehall	1,58	playful	happy	friendly	L
Looped	Kiasmos	Deep-House	1,51	relaxing	modern	contempo-rary	L
Radio Days	Fred Falke	Disco	1,68	retro	happy	cool	L
Essarai	Cantoma	Downbeat	1,56	chilled	relaxing	soft	L
Today Is The Same Day As Yesterday, But Yesterday Is Not Today	Noiserv	Dream Pop	1,64	soft	relaxing	warm	L
Dumptruck	Blokh4ed	Drum&Bass	1,57	modern	futuristic	young	L
Elle	Sophie	Dubstep	1,27	futuristic	modern	young	L
Bloop Bleep	Gary McFarland And His Orchestra	Easy-Listening	1,57	chilled	soft	warm	L
Summer Won't Wait (Extended Mix)	Marley James	EDM	1,97	young	modern	happy	L
Ghost Orchid	Benjamin Fröhlich	Electro	1,58	futuristic	modern	young	L
Weekend	Class Actress	Electro-Pop	1,58	young	modern	contempo-rary	L
Hungry	American Royalty	Electro-Rock	1,66	modern	young	cool	L
L'albero Del Pane	Gianmaria Testa	Flamenco	1,65	relaxing	warm	soft	L
Don't Joke With A Hungry Man (Part 2)	Spanky Wilson & The Quantic Soul Orchestra	Funk	1,51	happy	authentic	urban	L
One Of A Kind – Part One	Bill Bruford	Fusion-Jazz	1,52	creative	playful	retro	L
PMW (All I Really Need)	A$AP Rocky	Hip-Hop	1,64	urban	young	modern	L
Weapon Of Choice (Fish Go Deep Remix)	Fish Go Deep	House	1,57	young	modern	cool	L
Constellations	Darwin Deez	Indie-Dance	1,54	happy	friendly	young	L
Check My Heart	The Pastels	Indie-Pop	1,58	friendly	young	soft	L
Rebel Heart	The Shelters	Indie-Rock	1,78	happy	young	friendly	L
Meh-Teh	Tussle	Krautrock	1,49	futuristic	modern	young	L
Viva Tirado	Fania All-Stars	Latin	1,71	happy	creative	bright	
Shooting Star	Miles Sanko	Northern-Soul	1,64	happy	friendly	cool	L
Jiinti	Mocky	Nu-Jazz	1,65	relaxing	chilled	soft	L
The Weapon (Album Version)	Rush	Progressive-Rock	1,75	young	modern	authentic	L

Day Tripper	Shockabilly	Punk	1,40	modern	creative	futuristic	L
Devil's Pie	D'Angelo	R&B	1,45	modern	urban	young	L
Brothers On The Slide	Cymande	Rare-Groove	1,62	cool	urban	happy	L
Message From A Black Man	The Heptones	Reggae	1,65	happy	friendly	playful	L
Visionary	Farruko	Reggaeton	1,80	young	modern	urban	L
Já Mandei Botar Dendê	Nilze Carvalho	Samba	1,73	happy	friendly	playful	L
Say You Love Me	Monophonics	Soul	1,80	passionate	chilled	authentic	L
Tango Del Diablo	Soledad	Tango	1,59	classic	unique	creative	L
House Music	Maelstrom	Tech-House	1,53	modern	young	futuristic	L
Little Miss Sunshine (Little Miss Rain)	Lee Hazelwood	Traditional-Folk	1,61	friendly	warm	relaxing	L
Riddim Box	NB Funky	UK-Funky	1,53	young	modern	contempo-rary	L
A Dream Of A Spider	Andrew Thomas	Ambient	1,51	relaxing	chilled	soft	S
Amulet (Begin Remix)(feat. Peter Coyle)	Pharao Black Magic	Balearic	1,44	futuristic	modern	contempo-rary	S
The Soul Drummers	Ray Barretto	Boogaloo	1,74	happy	friendly	playful	S
How Insensitive	A.C. Jobim & Sting	Bossa-Nova	1,60	soft	relaxing	chilled	S
Something Goes Right	SBTRKT	Broken-Beats	1,51	young	modern	urban	S
Calypso Queen	Calypso Rose	Calypso	1,81	happy	friendly	playful	S
La Javanaise	Juliette Gréco	Chanson	1,67	soft	relaxing	passionate	S
Age Of Self	Jono McCleery	Contemporary-Folk	1,58	relaxing	soft	warm	S
Pumpkin Belly	Tenor Saw	Dancehall	1,65	young	urban	modern	S
Liberating Soul	Tesla Boy	Disco	1,69	retro	young	modern	S
Bluebird	Beach House	Dream Pop	1,75	soft	relaxing	chilled	S
Morning BJ	Point Point	Dubstep	1,51	futuristic	modern	young	S
What's A Girl To Do	Fatima Yamaha	Electro	1,57	futuristic	modern	creative	S
Lady	Chromatics	Electro-Pop	1,63	young	modern	contempo-rary	S
Blues Fight	Black Strobe	Electro-Rock	1,64	modern	young	urban	S
Holy Ghost	Bar-Kays	Funk	1,98	authentic	warm	relaxing	S
Drop A Gem On 'Em	Mobb Deep	Hip-Hop	1,51	urban	young	modern	S
The Nile	Kisses	Indie-Dance	1,74	young	happy	friendly	S
What Doesn't Kill You	Jake Bugg	Indie-Rock	1,84	young	urban	modern	S
Las Lomas De New Jersey	Marc Ribot Y Los Cubanos Postizos	Latin	1,78	friendly	warm	authentic	S
My Brother's Keeper	Stratus	Nu-Jazz	1,79	relaxing	chilled	soft	S
Deutscher Girls	Adam And The Ants	Punk	1,84	young	happy	cool	S
I Remember Well	The New Birth	Rare-Groove	1,63	happy	friendly	retro	S
La Mano Encima	Plaza Francia	Tango	1,65	passionate	authentic	natural	S
That's Us – Wild Combination	Joel Gibb	Traditional-Folk	1,63	soft	relaxing	chilled	S

2.3 Design and Measures

2.3.1 Independent Variables

The experimental design consists of two variables which are manipulated. That is *assortment size* and number of attributes (from now on referred to as *informativeness*). Each of them has two conditions. For *assortment size*, there is a *small* and *large* condition, for *informativeness* there is a *low* and *high* condition. This leads to a 2x2 full factorial design.

The variable *assortment size* is very inconsistent throughout previous research regarding choice overload. Small choice sets start for example from 2 options (Chernev, 2005) and reach up to 60 options (Diehl & Poynor, 2010), whereas the large choice sets begin with 5 alternatives (Chernev, 2005; Gourville & Soman, 2005) and end with 300 alternatives (Diehl & Poynor, 2010). As there is such a wide range represented, the conditions for *assortment size* are therefore further subdivided. The limitations of prior research, which have been presented before, support this procedure. In the *small assortment size* condition, participants are presented either 5, 10, 15, 20 or 25 songs, the playlists in the *large assortment size* condition contain either 30, 35, 40, 45 or 50 songs, resulting in 10 conditions (leading to a 10x2 design). The grouping of *small* and *large* will further be used for explanation of the implementation structure and to assign the participants, to make sure that each of them will test playlist sizes from both conditions. However, in the analysis the grouping of small and large will not be important anymore, as the real assortment sizes (from five to 50) will be looked at directly.

As mentioned before, the manipulation of the variable *informativeness* is also made up on two levels. The songs of the playlists assigned to the *low informativeness* condition are presented with two attributes: their title and the artist. In the *high informativeness* condition, the songs are displayed with their title, artist, the musical style and the three highest-rated GMBI adjectives from Lepa et al. (2020), hence six attributes. The decision for two versus six attributes comes from a design of Greifeneder et al. (2010), where the authors did a pre-test to check for relevant differences in number of attributes for their experiment. Even if a significant difference between one and six attributes was found in this pre-test, it makes more sense to present the stimuli of this study with the title and artist (instead of just the title) in the *low informativeness* condition, as it comes closer to a realistic scenario (in playlists you generally find the pieces with this type of information).

Each of the participants tests two conditions, hence two playlists. They are presented one of the *large* and one of the *small* playlists, and they also see each of the *informativeness* conditions. The assignment of which level of *assortment size* is combined with which level of *informativeness* is randomized, so is the order of those combinations. This within-subjects design was applied from the music study of Scheibehenne et al. (2009), as it allows to attenuate the effect of individual differences between participants.

2.3.2 Control Variables

For a manipulation check, two items were generated based on the manipulation check from an experiment by Chernev (2003). The first item is included to examine the perceived variety of the choice set (“How much variety did the set of songs offer?”), whereas the second item is intended to evaluate the perceived amount of given information (“How did you perceive the amount of given information to each song?”). Both questions should be answered on 5-point scales with the following options: “1 – overwhelming”, “2 – rather extensive”, “3 – adequate”, "4 – somewhat narrow" and "5 – very limited". The manipulation check is included to see how the manipulated variables are seen by the participants and to check how they affect the dependent variables. This check is especially important because the perceived variety is a function of assortment size (Broniarczyk et al., 1998; Van Herpen & Pieters, 2002) and as assortment size is one of the manipulated variables, the evaluation of the perceived variety is necessary to see whether the manipulation was successful.

To ensure that the attractiveness of the options in the playlists is held constant and does thus not influence the outcome (Bollen et al., 2010; Chernev & Hamilton, 2009; Goodman & Malkoc, 2012), two more items on the perceived quality of the playlists are included from Hadar & Sood (2014). Those are "How would you rate the overall quality of the song options you were presented with?" (*perceived quality*) and "Overall, how positively do you feel about the song options you were presented with?" (*positivity with options*). The evaluation of these items follows on a 7-point scale beginning with "1 – not at all positive" and ending with "7 – very positive".

Additionally, participants are asked to indicate the number of artists and songs they already knew from the playlists (in integer values on a slider starting from zero going up to ten) and whether they chose one of them. The addition of these questions seems necessary in order to consider or exclude a possible bias of the participants.

2.3.3 Dependent Variables

CHOICE-MAKING PROCESS. In order to evaluate the choice-making process of the participants, a series of questions regarding *enjoyment* ("How much did you enjoy making the choice?"), *difficulty* ("Did you find it difficult to make your choice of which song to pick?") and *frustration* ("How frustrated did you feel when making the choice?") during the choice-making process and *anticipated satisfaction* with the choice ("How satisfied do you think you will be if you listen to this song?") were included from a design by Iyengar & Lepper (2000). The rating scale to those items starts from "1 – not at all" and goes up to "7 – extremely". These questions were chosen as they allow a separate evaluation of the choice-making process, as participants will be – when these questions are asked – not yet be influenced by the outcome of their choice. They also prove to be good measures to examine choice overload, as satisfaction, decision difficulty and frustration are strongly linked to it (Chernev et al., 2015) and they were used in other experiments regarding choice overload (Bollen et al., 2010; Diehl & Poynor, 2010; Griffin & Broniarczyk, 2010; Reutskaja & Hogarth, 2009) as well. Furthermore, the decision for variables measuring the subjective state of the participants was made, as they include an ordinal scaling and give a more nuanced insight in the experience of choice overload, compared to the behavioral outcome. Measures like *choice deferral* for example are not suitable, as the experience with the choice is object of the investigation, hence participants have to make a choice.

As *switching likelihood* is another factor comprised by choice overload (Chernev, 2003; Chernev et al., 2015), another item is added to the set of questions regarding the choice-making process: “If you had the chance to change your decision, how likely is it that you would do it?”, the scaling to it is the same as for the previous items.

MUSIC LISTENING EXPERIENCE. To measure music listening experience (the outcome of the choice), a set of items was generated from previous research as well. Three questions are from an experiment by Anglada-Tort et al. (2018) and serve to collect data about the *aesthetic properties* of the stimuli. They include how much the participants *liked* the music (“How much do you like this song?”, with a scale from “1 – dislike strongly” to “7 – like strongly”), the perception of the *emotional expressivity* (“How would you rate the emotional expressivity of this song?”) and *musical quality* (“How would you rate the musical quality of this song?”). The last two questions are also answered on 7-point scales starting at “1 – very bad” and end at “7 – very good”.

At further questions, participants have to evaluate the excerpts in terms of *subjective value*. That is the degree of how likely a) they think the song could be *commercially successful* (“How likely do you think this song would succeed commercially?”), b) they would go to a *concert* (“How likely would you go to a concert of the artist?”) and c) it is that they *recommend it* (“How likely is it that you would recommend this song to a friend?”) (Anglada-Tort et al., 2018). The rating of these questions follows on a 7-point scale as well (from “1 – very unlikely” to “7 – very likely”). In order to evaluate the music listening experience more specifically, there is a need for measures that have not been adapted from prior research on choice overload, which is why those items are included.

In accordance with the items on enjoyment with the choice-making process and the predicted satisfaction, participants are also going to be asked how much they *enjoyed* the sample they have listened to (*post-choice enjoyment* - “How much did you enjoy the sample you listened to?”) and how much they are *satisfied* with their choice (*post-choice satisfaction* - “How satisfied were you with the song you listened to?”). These items come from the same experiment from Iyengar & Lepper (2000) as the questions evaluating the choice-making process, and have to be answered on the same scale. The difference is that they are intended to examine the outcome of participants' choice and not the choice-making process itself. In this way, the effects can be viewed separately from one another.

When the expectation and the final perception of a product do not meet, people experience disconfirmation, this can be either positive or negative (Diehl & Poynor, 2010). This fact influences outcome satisfaction and might thus have an impact on the listening experience. Therefore another item to measure *expectation disconfirmation* from Diehl & Poynor (2010) is included ("Did the song you chose meet your expectations?") with the original 9-point scale from "1 – much worse than I expected" to "9 – much better than I expected".

At the last question, participants have to indicate in integer values on a slider how they would rate their *overall listening experience* from "0 – very poor" to "100 – very good". This element was added to complete the dependent measures and to give the participants the opportunity to summarize their experience.

The chosen outcome variables require a forced-choice paradigm, that means that answers to each question are obligatory and that questions may not be skipped.

2.3.4 Moderator Variables

According to Schwartz et al. (2002), the tendency of individuals to satisfice or maximize has an impact on satisfaction and regret in decision processes. In order to evaluate participants in terms of their behavior in this regard, items from the *Maximization Scale* developed by Schwartz et al. (2002) are included as a possible moderator (*propensity to maximize*) in this design. The items are chosen based on their correlation values with the maximizing scale. Five questions with the highest values are taken in: "When I watch TV, I channel surf, often scanning through the available options even while attempting to watch one program.", "When I am in the car listening to the radio, I often check other stations to see if something better is playing, even if I'm relatively satisfied with what I'm listening to.", "No matter how satisfied I am with my job, it's only right for me to be on the lookout for better opportunities.", "I often fantasize about living in ways that are quite different from my actual life." and "Streaming videos is really difficult. I'm always struggling to pick the best one.". The last item was modified, as it was originally about renting videos and not streaming videos, and renting videos is no longer up to date. All items have to be rated on a 7-point scale starting from "1 – completely disagree" and ending at "7 – completely agree".

2.4 Procedure

2.4.1 Implementation

To be able to execute this design, an online survey is implemented using the survey application *LimeSurvey*[6], Version 3.16.5. It brings the advantage, that no software is needed, and that people can participate using a link which directs them to the survey. The collected data is saved on a server and is available online as well. The execution of the survey and the evaluation of the data is thus not bound to one place. LimeSurvey offers a wide range of functions. Many of them can be implemented using the application's user interface, adjustments can be made using JavaScript, CSS and LimeSurvey's own *Expression Manager*.

To create a question, it is necessary to know which question type can best be used for it. The question type, as well as the answer options can be easily adjusted using the LimeSurvey user interface. There are several question types in the present survey. The *Array* type is used for all questions to be answered on Likert-type scales. This type is suitable because *one* rating scale can be used for more questions and the different answer options are presented horizontally. For simple questions with two to four answer options, the *List* type is used, since this alternative offers the clearest presentation of short items. The *Numerical Input* type is selected for age specification and the questions which include a slider to answer. To define the variables for randomization purposes and a preselection of answer options, the *Equation* type is used. This type allows to set equations and conditions. Questions can be put together in question groups which are displayed separately, to make the overview of the survey clearer. This survey comprises eleven question groups and a total of 40 questions, some of which are hidden.

RANDOMIZATION. As the organization of an assortment influences the perceived variety of it, with disorganized assortments perceived as offering more variety (Kahn & Wansink, 2004) it is important to keep the organization across the assortments constant by randomizing the order of the stimuli. This excludes a possible effect of assortment organization on the outcome. In order to make a randomization of the condition combinations possible, variables had to be defined. They are shown in Table 2. To implement the randomization, four questions of the type *Equation* were created, one of them for each variable. The questions have the same name as the variables and a condition was set so that these variables get random values assigned either from one to five (*randnumber* and *randnumber2*), or from one and two (*OrderInfo* and

[6] https://www.limesurvey.org/

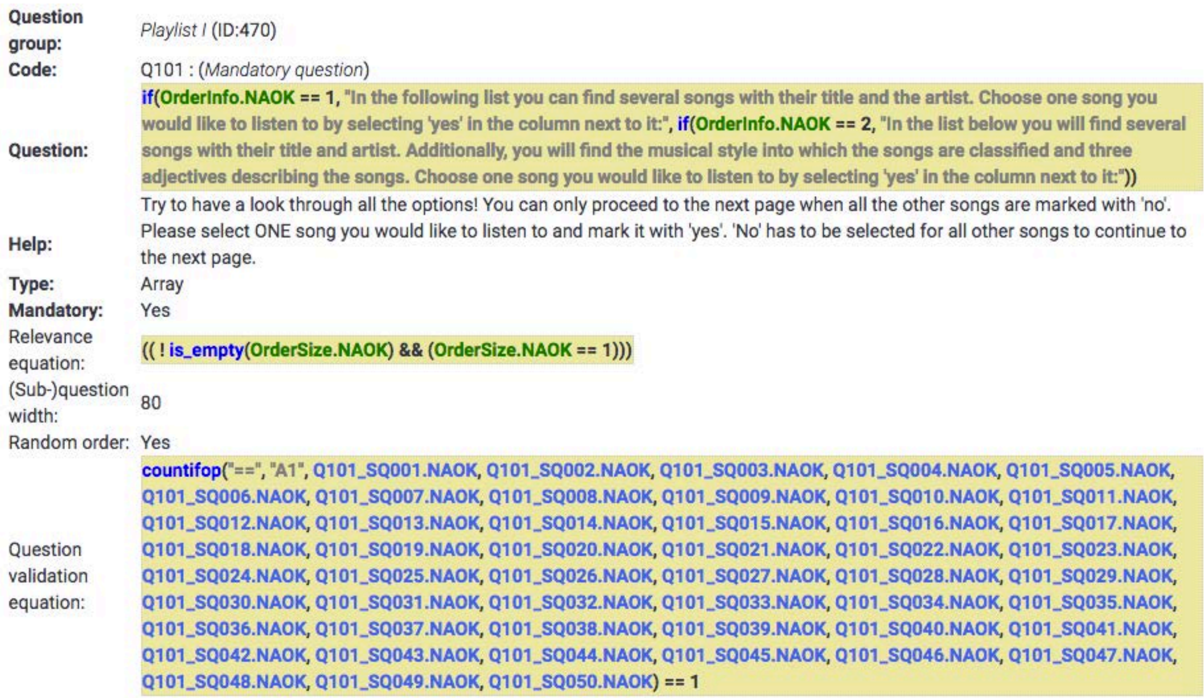
OrderSize). The values for the variables are set before the first condition starts, so that it is predefined whether a participant will see the small or the large pool of songs first, the order of the informativeness conditions, and the number of song options presented in each pool.

Table 2. *Randomization variables and their values.*

randnumber	randnumber2	OrderInfo	OrderSize
presented options in small playlist	*presented options in large playlist*	*order of playlist informativeness*	*order of playlist size*
1: 5 options	1: 30 options	1: low/high	1: small/large
2: 10 options	2: 35 options	2: high/low	2: large/small
3: 15 options	3: 40 options		
4: 20 options	4: 45 options		
5: 25 options	5: 50 options		

PLAYLISTS. In the beginning, the playlists were created with all possible songs and conditions as an *Array* type. This allows a good presentation. Depending on the value of the randomization variables, conditions and relevance equations determine which songs from this entire pool are presented with which level of informativeness. The song options in the playlist are presented in randomized order as well. The answer options to each song are “yes” (A1) and “no” (A2). To make sure that participants pick just *one* song from each playlist, a question validation equation was set (see Figure 2). It makes it impossible for participants to proceed to the next page if more than one song is marked with “yes”. An example of a question containing a possible first playlist is displayed in Figure 2. This would be a *small* playlist, as it will just be shown if *OrderSize* = 1. The question text specifies which text is presented, depending on how *OrderInfo* is selected. The selection of songs is done in the subquestions of each playlist question. As an answer is mandatory, all the other songs must be answered with “no”. The biggest pool consists of 50 songs, so a hidden *Equation* question named *Preselectorz* was implemented, which preselects the answer “no” for every song. This was done to save participants’ time and effort. With the further use of relevance equations, the actual excerpt of the selected song can be streamed in one of the following question groups.

The definition of the variables and the setting of the conditions were done with the *Expression Manager* of LimeSurvey. HTML commands were used to modify the graphical presentation of the playlists. The whole survey can be found as a *.lss*-File on the disk. In the next section, the structure of the survey will be explained in detail.



Question group:	*Playlist I* (ID:470)
Code:	Q101 : (*Mandatory question*)
Question:	if(OrderInfo.NAOK == 1, "In the following list you can find several songs with their title and the artist. Choose one song you would like to listen to by selecting 'yes' in the column next to it:", if(OrderInfo.NAOK == 2, "In the list below you will find several songs with their title and artist. Additionally, you will find the musical style into which the songs are classified and three adjectives describing the songs. Choose one song you would like to listen to by selecting 'yes' in the column next to it:"))
Help:	Try to have a look through all the options! You can only proceed to the next page when all the other songs are marked with 'no'. Please select ONE song you would like to listen to and mark it with 'yes'. 'No' has to be selected for all other songs to continue to the next page.
Type:	Array
Mandatory:	Yes
Relevance equation:	((! is_empty(OrderSize.NAOK) && (OrderSize.NAOK == 1)))
(Sub-)question width:	80
Random order:	Yes
Question validation equation:	countifop("==", "A1", Q101_SQ001.NAOK, Q101_SQ002.NAOK, Q101_SQ003.NAOK, Q101_SQ004.NAOK, Q101_SQ005.NAOK, Q101_SQ006.NAOK, Q101_SQ007.NAOK, Q101_SQ008.NAOK, Q101_SQ009.NAOK, Q101_SQ010.NAOK, Q101_SQ011.NAOK, Q101_SQ012.NAOK, Q101_SQ013.NAOK, Q101_SQ014.NAOK, Q101_SQ015.NAOK, Q101_SQ016.NAOK, Q101_SQ017.NAOK, Q101_SQ018.NAOK, Q101_SQ019.NAOK, Q101_SQ020.NAOK, Q101_SQ021.NAOK, Q101_SQ022.NAOK, Q101_SQ023.NAOK, Q101_SQ024.NAOK, Q101_SQ025.NAOK, Q101_SQ026.NAOK, Q101_SQ027.NAOK, Q101_SQ028.NAOK, Q101_SQ029.NAOK, Q101_SQ030.NAOK, Q101_SQ031.NAOK, Q101_SQ032.NAOK, Q101_SQ033.NAOK, Q101_SQ034.NAOK, Q101_SQ035.NAOK, Q101_SQ036.NAOK, Q101_SQ037.NAOK, Q101_SQ038.NAOK, Q101_SQ039.NAOK, Q101_SQ040.NAOK, Q101_SQ041.NAOK, Q101_SQ042.NAOK, Q101_SQ043.NAOK, Q101_SQ044.NAOK, Q101_SQ045.NAOK, Q101_SQ046.NAOK, Q101_SQ047.NAOK, Q101_SQ048.NAOK, Q101_SQ049.NAOK, Q101_SQ050.NAOK) == 1

Figure 2. *Screenshot from the LimeSurvey user interface of a question from the question group 'Playlist I'.*

2.4.2 Execution

The experiment is conducted as an online survey. This choice offers the opportunity, that participants will be in their natural environment, where they normally listen to music. This increases the external validity of the experiment. LimeSurvey provides a link to the survey that can be sent to potential participants.

With the reception and opening of the LimeSurvey link, an introduction page is shown. It includes a short summary of the task and a remark that it is recommended to do the survey in a quiet environment using headphones. It also states that the responses are anonymous and that participants have to be 18 years or older in order to participate. In the following pages of the survey, the questions are presented in question groups.

The first question group displays a consent page. It informs about the task and the duration (15-20 minutes), as well as data protection. Contact details of the author can be seen as well. To proceed, participants have to agree that they have read the information, voluntarily participate and are 18 years of age or older. If they disagree, the survey stops at this point.

Demographic information is requested in the second question group. Participants have to enter their age as an integer value and can specify their gender. Not only "female" and "male" are the available options. The option "Other" with a text field for further specification is also available, as well as "I prefer not to say". The survey can only be conducted in English. However, since non-native speakers also take part in the survey, the English level of the respondents is recorded to see whether any comprehension problems have an effect on the results.

The following question group contains more detailed instructions for the upcoming procedure. It describes the structure of the survey and explicitly states that participation should take place with a stable internet connection (as the excerpts will be streamed), in a quiet environment and with headphones. It also includes a remark that the execution is possible on a mobile device, but that it is recommended to use laptop or desktop system.

On the next page, participants are asked to adjust their volume. By clicking “Play” on a small media player, they can listen to an automated male voice giving them instructions to adjust their volume on their device. The audio file was normalized in loudness as well as the stimuli (see section 2.2). This is important, as the stimuli should be listened to with an equal volume, which should not be changed whilst listening. This question group also includes the randomization of the variables described in section 2.4.1, but they are hidden from the participants.

In the following question group (*Playlist I*), the first playlist is presented. The size and informativeness depends on the selected randomization variables. The following scenario is presented at the beginning of the page: “Imagine you are at home and you would like to listen to new music. On your personal music player, there is a list with songs recommended for you:”. This is followed by a description of the information given to each song. Participants should try to have a look through all the options and select *one* song they would like to listen to. An example of an excerpt from a playlist with the *high informativeness* condition is shown in Figure 3. Only if one song is selected with “yes” and all the other songs are marked as “no”, participants are allowed to proceed to the next page. This means that participants are forced to make a choice and are not allowed to defer their choice.

The next question group (*Choice I*) comprises of the items recording the choice-making process (see section 2.3.3) and the control variables (see section 2.3.2). The separation of choosing and the evaluation of this choosing process from the actual listening allows an independent evaluation of participants affective responses in the selection process. In three *Array* type questions, people are asked about the choice-making process, the perceived variety and amount of information, and the quality of the choice set. Each array has its own rating scale. At the bottom, there is one item asking how many of the titles or artists were already known to the participants. The answer must be indicated on a slider from zero to ten.

	yes	no
Devil's Pie – D'Angelo R&B ***modern - urban - young***	○	●
Meh-Teh – Tussle Krautrock ***futuristic - modern - young***	○	●
Day Tripper – Shockabilly Punk ***modern - creative - futuristic***	○	●
Já Mandei Botar Dendê – Nilze Carvalho Samba ***happy - friendly - playful***	○	●
Ghost Orchid – Benjamin Fröhlich Electro ***futuristic - modern - young***	○	●

Figure 3. *Excerpt from a playlist with high informativeness.*

The forthcoming page displays the question group *Listening I*. Right at the beginning is a small media player, which plays the 2-minute excerpt of the selected song from the first playlist. Above it, the title and artist are displayed. The instructions say that participants should use headphones and listen first, before continuing with the further questions. After listening, there is a space next to the media player that must be ticked in order to proceed. The first question asks if participants chose a song they already know. What follows are the questions examining the music listening experience (see section 2.3.3). They are made up of five *Array* type questions consisting of the aesthetic properties, subjective value, satisfaction and expectation. Again, each array has its own rating scale. For the last part of this question group, participants have to rate their overall listening experience on a slider from zero to 100.

From this point on, the procedure will repeat and start with a new playlist in the condition the participants have not yet seen (*Playlist II, Choice II* and *Listening II*).

The last question group contains the five selected elements of the *Maximization Scale*. These are also arranged in an array since they should all be rated on the same rating scale. It seemed important to place these items at the end of the survey, as this could avoid influencing the participants in their responses due to the unambiguity of these questions (they could give the participants a hint what the research object is).

On the end page of the survey, people are thanked for participating. It also includes the remark that students who are enrolled in M.Sc. Audiocommunication and –technology get 30 credited minutes for participating. Except for the question about participants gender, *all* questions must be answered. This means that participants cannot skip a question or abstain from answering.

2.5 Data Analysis

The data will be analyzed using the software *IBM SPSS Statistics 23.0 Desktop*[7]. As LimeSurvey offers an export function of the results of a survey to a SPSS file, the consideration of other possibilities to analyze the data was therefore not necessary.

This chapter describes the further procedure with the data. First, we will have a look at how the randomization of the different conditions turned out, by checking how often each condition was assigned to a participant. Second, we are going to check if certain items can be merged into single scales, by examining the internal consistency. Furthermore, it will be examined if the level of English has an influence on the outcome.

2.5.1 Frequencies of Conditions

To check how LimeSurvey assigned the participants to each condition by selecting the randomization variables described in section 2.4.1, a look at the frequencies of each condition is necessary. Using a 10x2 experimental design (*assortment size x informativeness*) results in 20 conditions. With approximately 100 people participating, each of them testing two conditions, we get a total of 200 cases. If the distribution of the assignments was equal, each condition would have been tested ten times.

By looking at Table 3, one can see that the distribution among the conditions is not equal. The middle sized assortments were assigned more frequently, whereas a decrease is observable to the small and large assortments, reminding of a normal distribution. Nevertheless, there are still enough cases for each of the condition combinations.

Table 3. *Absolute frequencies of the condition assignments.*

assortment size	**5**	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	**50**	**total**
low informativeness	11	6	10	15	13	11	11	8	8	8	***101***
high informativeness	6	7	16	8	9	17	7	11	10	10	***101***
total	**17**	**13**	**26**	**23**	**22**	**28**	**18**	**19**	**18**	**18**	***202***

[7] https://www.ibm.com/support/pages/downloading-ibm-spss-statistics-23

2.5.2 *Internal Consistency Reliability*

In order to check whether items intended to measure *one* construct have internal consistency, Cronbach's α will be calculated. If the value is suitable, the items will be aggregated into *one* scale/measure by calculating the mean from these items for each participant.

CHOICE-MAKING PROCESS. The first construct is intended to measure *process satisfaction*, hence positive effects when making choices. This construct constitutes of *enjoyment* and *anticipated satisfaction* ($\alpha = .70$). The *negative effects* of the choice-making process are composed of *difficulty* and *frustration* ($\alpha = .70$). *Switching likelihood* can be regarded as a negative effect as well, but by including this item the α-value would drop to .66. It will therefore be handled as a single measure in the further analysis. As a scale for choice overload is desired, it is tested whether all of the positive and negative effects can be merged into one *choice overload* measure. For doing this, the items *enjoyment* and *anticipated satisfaction* have to be transformed into new variables with inversed rating. Taking the *negative effects*, *switching likelihood* and inversed *process satisfaction* together, a value of .70 for *choice overload* is reached (by excluding *switching likelihood* the α-value remains equal).

MUSIC LISTENING EXPERIENCE. In line with the results from Anglada-Tort et al., (2018), the items measuring the *aesthetic properties* ($\alpha = .85$) and the *subjective value* ($\alpha = .75$) of the music can be merged into single measures as well. *Outcome satisfaction* is made up of *post-choice satisfaction* and *post-choice enjoyment* ($\alpha = .96$). As *expectation disconfirmation* has a different rating scale than all the other items, it is kept as a single item. Performing an exploratory factor analysis (extraction method: maximum likelihood, rotation method: Varimax with Kaiser normalization) revealed high factor loadings for all items intended to measure *music listening experience* in one factor. Following these results, a *listening experience* scale is developed by merging all items (except *expectation disconfirmation*) which have been mentioned above ($\alpha = .93$).

MODERATOR. Five items from the maximization scale from Schwartz et al. (2002) were included in the experiment to measure participant's *propensity to maximize*. Cronbach's α for these five items is .61, by deleting the last item (*streaming videos*) the value increases to .63. The low α-value might be a result of reducing the items from the original maximization scale and not choosing the "right" items. As this value cannot be increased by deleting further items, an inter-item correlation is processed. The correlation of the left four items reveals that these items are nevertheless related ($r > .15$, $p < .01$, for all correlations), and can therefore be merged into a scale.

CONTROL VARIABLES. All of the items to check the manipulations or to control a possible influence of other variables are single items and are not part of a construct. Only the items *perceived quality* and *positivity with options* are meant to measure the construct *option attractiveness* ($\alpha = .86$).

2.5.3 Influence of English Level

As stated before, the online survey was only available in English. With nearly all the participants being German with different levels of the English language, there is a need to examine a possible effect of it on the outcome. This could be because of translation or comprehension problems (individuals with a lower English level might understand question and items differently than native speakers and could therefore answer totally different, even though they have been assigned to the same condition). To examine whether the English level has an effect, a comparison of mean values is performed in form of an ANOVA with all the outcome and moderator variables described in chapter 2.5.2 over the factor *English level*.

Results show that there is no significant difference between English levels for the measures *expectation disconfirmation* ($F(4, 197) = 1.19$, $p = .32$), *choice overload* ($F(4, 197) = .22$, $p = .93$) and *listening experience* ($F(4, 197) = 2.23$, $p = .07$). However, the situation is different for the *propensity to maximize* scale ($F(4, 197) = 5.21$, $p = .001$). By looking at the mean values for each level displayed in Table 4, it is clear that *first language* depicts an outlier, being considerably higher than the other groups and the total mean. Multiple comparisons in form of a Hochberg's GT2 post hoc test reveal significant differences between *first language* and *proficient*, *beginner*, *intermediate*. Hochberg's GT2 was used because of homogeneity of variances (Levene's test showed that equal variances can be assumed, $p = .2$) and a considerable difference in the number of cases (Field, 2013).

Even if the differences between *first language* and the other groups are significant for this measure, it is assumed that that it is not due to the language level, since in the other measures no differences could be detected. The results could also be due to the small number of participants who indicated their English level as *first language.* These individuals could all have a similarly high *propensity to maximize.* Thus, it is assumed that the English level of participants has no influence on the results.

Table 4. *Mean, standard deviation and frequencies of the variable propensity to maximize.*

English level	*mean*	*standard deviation*	*N*
first language	5.2	1.1	5
proficient	3.5	1.2	51
intermediate	3.7	1.3	38
beginner	3.0	1.1	6
none	3.3	.0	1
total	*3.6*	*1.3*	*101*

As the experimental items have been aggregated into meaningful constructs and the influence of the English level could be excluded, the results can be looked at in the following chapter.

3 Results

3.1 Control Variables

To ensure that the manipulation of the independent variables was successful across participants, two items were included in the experimental design. The results will be presented in this section. Moreover, two variables which have not been controlled, but might have an influence on the outcome, have to be checked as well. That is *option attractiveness* and *dominant option*. Whether they have an impact is going to be examined as well.

MANIPULATION CHECK. Both items for the manipulation check of the independent variables *assortment size* and *informativeness* had rating scales starting from "1 – overwhelming" to "5 – very limited". As these items are questions which ask about the perceived variety and amount of given information of the options in the choice sets, the scale can be confusing and is therefore inverted. This means that a low perceived variety and amount of given information also have a low value and vice versa.

The songs in the low *informativeness* condition were perceived having a significantly smaller amount of information ($M = 1.77$, $SD = .86$) than the songs in the high *informativeness* condition ($M = 3.13$, $SD = .88$), $t(200) = -11.09$, $p < .001$.

The perceived variety differed statistically significant over the different assortment sizes as well, $F(9, 192) = 5.96$, $p < .001$. As the results of an ANOVA do not provide information about the relationship between the two variables, a correlation is performed to see how *assortment size* influences *perceived variety*. The results demonstrate a moderate positive correlation between *assortment size* and *perceived variety* ($r = .42$, $p < .001$), suggesting that an increase in *assortment size* rises the *perceived variety* as well.

The results indicate that the manipulation of the independent variables was successful.

OPTION ATTRACTIVENESS. To ensure that the perceived quality of the playlists does not influence the outcome, an ANOVA was performed. The results show that there is no statistically significant difference between assortment sizes in terms of *option attractiveness*, $F(9, 192) = .95$, $p = .48$. This means that the playlists were perceived as having a similar quality and a possible impact of *option attractiveness* on the outcome can be excluded. This leads to an exclusion of hypothesis 4.

DOMINANT OPTION. Two items were included to check whether the existence of a dominant option (the amount of songs known to the participants and if they chose one of them) influences the rating of the dependent variables.

The results of an ANOVA reveal that the values of the outcome variables *expectation disconfirmation* ($F(7, 194) = .66$, $p = .71$), *choice overload* ($F(7, 194) = 1.34$, $p = .23$) and *listening experience* ($F(7, 194) = 1.42$, $p = .18$) do not differ statistically significant across the *number of known songs* ($M = 1.37$, $SD = 1.6$). This suggests that if a certain amount of songs has already been known to the participants, it did not affect their choice-making process and the outcome of the choice.

However, there seems to be difference in the outcomes when people chose a song they already knew before ($N = 27$). Participants who chose a song they already knew reported significantly lower levels of *choice overload* ($M = 3.05$, $SD = .88$) compared to participants who did not ($M = 3.62$, $SD = 1.13$), $t(200) = -2.49$, $p = .014$. Furthermore, the rating of *listening experience* was significantly higher when the choice included an already know song ($M = 5.63$, $SD = 1.23$) rather than an unknown ($M = 4.19$, $SD = 1.29$), $t(200) = 5.42$, $p < .001$. Strangely, participants rated the song they already knew "better than expected" ($M = 6.52$, $SD = 1.86$) compared to the other group ($M = 5.48$, $SD = 2.23$), $t(200) = 2.23$, $p = .023$.

The results suggest that an influence of a *dominant option* could not be controlled by selecting relatively unknown songs as stimuli. It is also clear that when participants selected a song they already knew, the rating of the dependent variables differs from that of the people who did not know the song they chose. It is therefore recommended to include the variable *dominant option* (in form of whether participants chose a song they already knew or not) in the further analysis to see how it influences the outcomes across *assortment size* and *informativeness*.

3.2 Choice-Making Process

This section deals with the impact of *assortment size* and *informativeness* on the choice-making process of the participants. Additionally, the effect of *propensity to maximize*, as well as the choice of a *dominant option* will be evaluated.

3.2.1 The Impact of Assortment Size on the Choice-Making Process

PROCESS SATISFACTION. First, it is examined how *process satisfaction* is influenced by the size of an assortment. An ANOVA shows that there are no statistically significant differences in *process satisfaction* levels across *assortment size* ($F(9, 192) = 1.41$, $p = .19$). A correlation only shows a marginally significant negative linear relationship ($r = -.13$, $p = .054$). Analogous to the results of Reutskaja & Hogarth (2009), a quadratic curve fitting is performed to see if there is a curvilinear relationship between *process satisfaction* and *assortment size*. The results demonstrate no significance and a very low explanation of variance ($R^2 = .018$, $F(2, 199) = 1.80$, $p = .17$). A linear regression reveals no significant results as well. However, by looking at the mean values in Figure 4, a peak in *process satisfaction* levels is observable at 25 and 30 options ($M = 4.41$, $SD = 1.29$), whereas a minimum is present at 40 options ($M = 3.45$, $SD = 1.09$).

NEGATIVE EFFECTS. The results of an ANOVA reveal that the rating of *negative effects* across *assortment size* differs statistically significant ($F(9, 192) = 2.72$, $p = .05$). The mean values across *assortment size* (Figure 4) suggest that a linear relationship is possible, as after a minimum is reached at 15 options ($M = 2.85$, $SD = 1.21$), the *negative effects* increase with an increase in *assortment size* ($r = .26$, $p < .001$). A linear regression with *negative effects* as the dependent variable and *assortment size* as the independent variable is significant, whereas only 7 % of the variance of *negative effects* can be explained by the factor *assortment size* ($R^2 = .069$, $F(1, 200) = 14.86$, $p < .001$). The estimated increase in *negative effects* is .03 with an addition of each 5 options to the assortment ($\beta = .03$, $t(200) = 3.86$, $p < .001$).

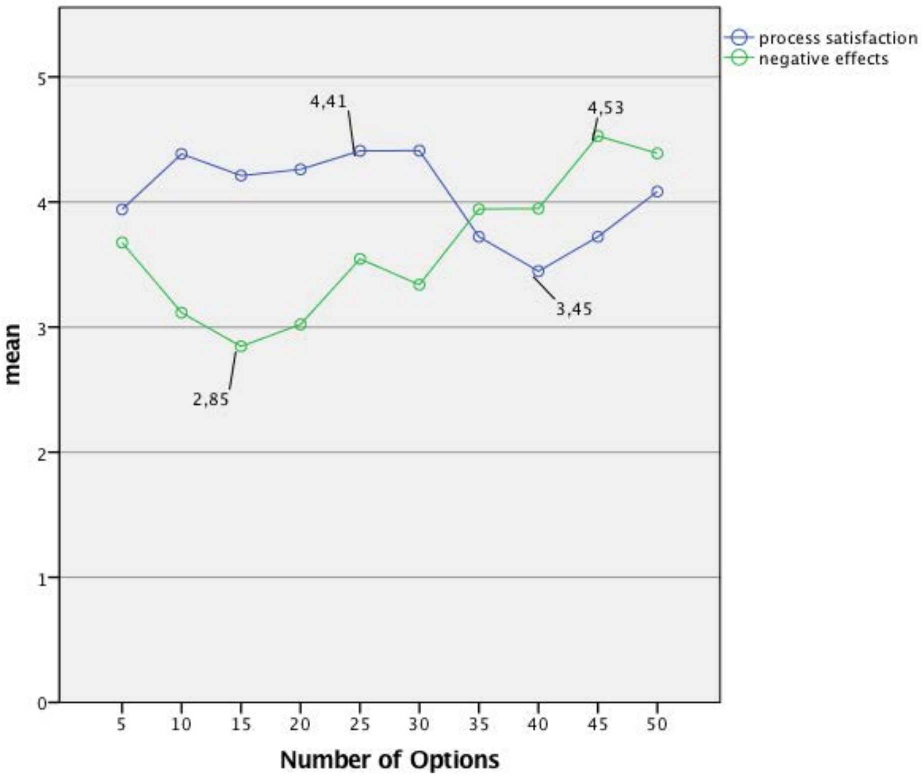


Figure 4. *Mean values for process satisfaction and negative effects across assortment size.*

Figure 4 shows that *process satisfaction* is low in the smallest assortment and reaches a plateau across 10 to 30 options. From this point on it decreases until it reaches a minimum at 40 options. An increase in *assortment size* also leads to an increase in process satisfaction from 40 up to 50 options. The course of *negative effects* is reversed. The level decreases from 5 to 15 options. A further increase in *assortment size*, from 15 options up, increases *negative effects* as well until the maximum is reached at 45 options. At 50 options a slight decrease is observable. Striking is that the curves of *process satisfaction* and *negative effects* cross between 30 and 35 options. Before this point, *process satisfaction* is higher, whereas after this point it is lower than *negative effects*. The largest difference of means is at 15 options ($MD = 1{,}36$).

SWITCHING LIKELIHOOD. As *switching likelihood* decreases the reliability of *negative effects* when included, it is therefore looked at separately. Examining *switching likelihood* as a function of *assortment size*, no significant effect can be found ($F(9, 192) = 1.05$, $p = .39$). However, a correlation of *assortment size* and *switching likelihood* reveals a statistically significant linear relationship ($r = .20$, $p < .001$).

CHOICE OVERLOAD. As all of the mentioned outcomes show a linearity over *assortment size*, they are aggregated into one *choice overload* measure, which has already been described in section 2.5.2. This procedure might make it possible to get significant results for all scales intended to measure choice overload, by merging them. As *process satisfaction* has been inverted before, a high value means a high level of *choice overload*.

The results of an ANOVA show that there are highly significant differences of *choice overload* levels across *assortment size* ($F(9, 192) = 2.69$, $p = .006$). A correlation reveals a significant linear relationship between *choice overload* and *assortment size* ($r = .26$, $p < .001$). These results are also supported by performing a linear regression, showing a significant influence of *assortment size* on *choice overload*, of which only 7 % of the variance can be explained through *assortment size* ($R^2 = .068$, $F(1, 200) = 14.63$, $p < .001$). It is estimated that an increase of 5 options to an assortment increases the level of choice overload by .02 ($\beta = .02$, $t(200) = 3.82$, $p < .001$). These results are very similar to those with *negative effects* as the dependent variable. Figure 5 shows that choice overload is higher for smaller choice sets until it reaches a minimum at 15 options (at this size, the mean difference between *process satisfaction* and *negative effects* is also the largest). Afterwards it steadily increases. The graphical presentation of the mean values supports the idea of a linear relationship.

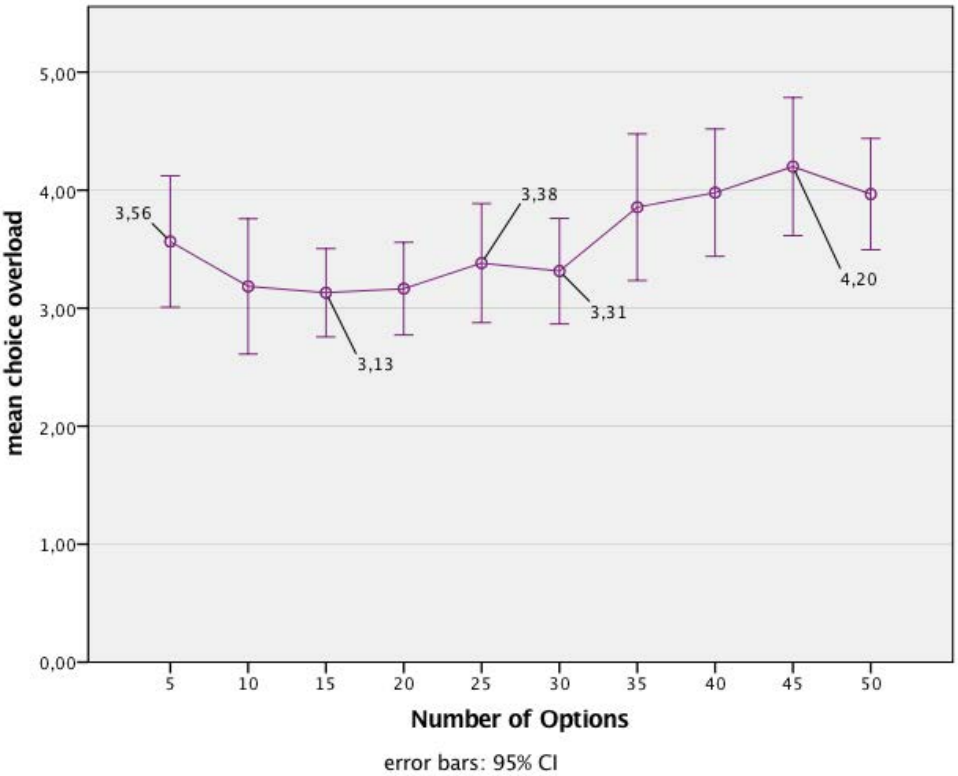


Figure 5. *Mean values for choice overload across assortment size.*

3.2.2 The Influence of Informativeness on the Relationship between Assortment Size and the Choice-Making Process

To see whether this linear relationship is also present when the number of given attributes changes, the outcome variables are examined with *informativeness* as an additional factor.

PROCESS SATISFACTION. When facing a playlist with high *informativeness*, participants reported significantly higher level of *process satisfaction* ($M = 4.36$, $SD = 1.28$) compared to playlists in the low *informativeness* condition ($M = 3.81$, $SD = 1.23$), $t(200) = -3.09$, $p = .002$. However, a significant difference of *process satisfaction* levels across assortment size can neither be found in the low *informativeness* condition ($F(9, 91) = .88$, $p = .55$) nor in the high *informativeness* condition ($F(9, 91) = 1.34$, $p = .23$). When presented with just two attributes, no linearity is existent between *process satisfaction* and *assortment size* ($r = -.02$, $p = .41$), whereas a significant linear relationship is observable when presented six attributes ($r = -.23$, $p = .009$). To examine this linearity in the high *informativeness* condition, a linear regression is implemented. The model shows a significant influence of *assortment size* on *process satisfaction*, which is able to explain 6 % of variance ($R^2 = .055$, $F(1, 99) = 5.76$, $p = .018$). The *process satisfaction* level decreases with every 5 options added to a choice set by -.02

(β = -.02, t(99) = -2.40, p = .018). Figure 6 visualizes the difference between *process satisfaction* in the low and high informativeness condition. It is always higher with a high number of attributes, except for one point (40 options), where *process satisfaction* levels of both conditions meet.

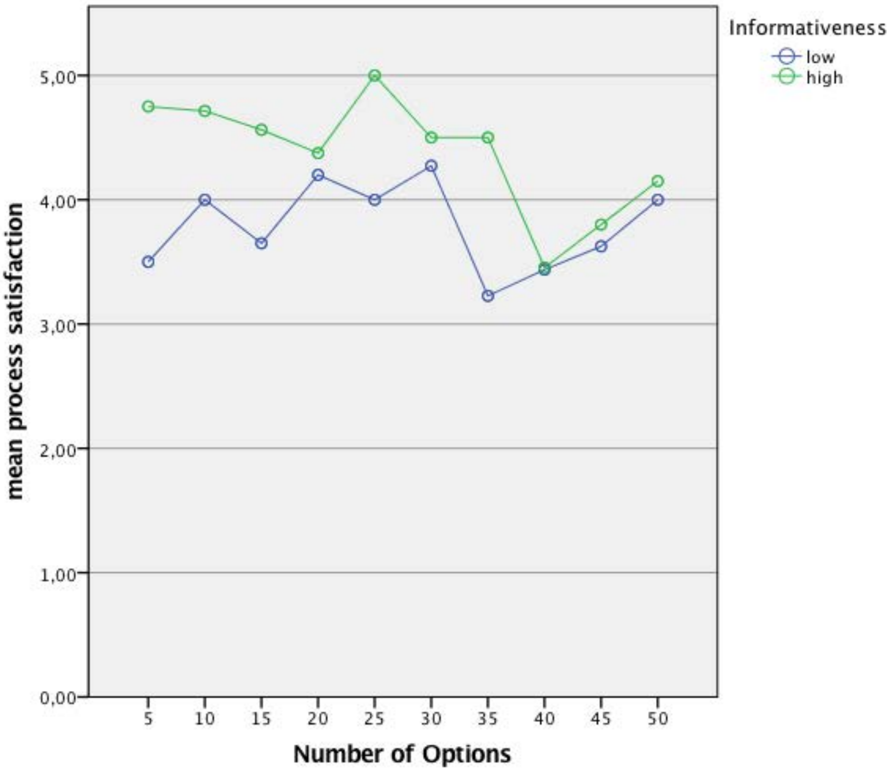


Figure 6. *Mean values for process satisfaction across assortment size for each informativeness condition.*

NEGATIVE EFFECTS. A significant difference between the two *informativeness* conditions for *negative effects* could not be detected, $t(200) = 1.15$, $p = .27$. The differences of *negative effects* levels across *assortment size* are for low as well as high *informativeness* non-significant ($F(9, 91) = 1.30$, $p = .25$; $F(9, 91) = 1.63$, $p = .12$, respectively). This is not the case when the relationship is looked at with no consideration of the *informativeness* condition. A significant linear relationship between *negative effects* and *assortment size* is nevertheless existent for both low ($r = .21$, $p = .017$) and high *informativeness* ($r = .33$, $p < .001$), with the linearity in the high *informativeness* condition being a bit higher. For both conditions, a significant linear model is applicable ($R^2 = .045$, $F(1, 99) = 4.67$, $p = .033$, $R^2 = .111$, $F(1, 99) = 12.42$, $p = .001$; low, high). They both predict an increase in *negative effects*, whereas in the high *informativeness* condition the increase is higher with an addition of 5 options each compared to low *informativeness* ($\beta = .026$, $t(99) = 2.16$, $p = .033$, $\beta = .037$, $t(99) = 3.52$, $p = .001$; low, high).

The relationship between process satisfaction and negative effects in each informativeness condition can be seen in Figure 10 and Figure 11 in the appendix.

SWITCHING LIKELIHOOD. There is no statistically significant difference between *switching likelihood* levels in the low and high *informativeness* condition, $t(200) = .53$, $p = .59$. An ANOVA shows no significant differences between *switching likelihood* ratings across *assortment size* for low and high *informativeness* ($F(9, 91) = .62$, $p = .78$; $F(9, 91) = 1.03$, $p = .420$, respectively). Nevertheless, the significant linear relationship is still existent when divided in low and high *informativeness* ($r = .18$, $p = .04$; $r = .24$, $p = .01$, respectively).

CHOICE OVERLOAD. Participants reported significantly higher levels of *choice overload* when choosing from playlists in the low *informativeness* ($M = 3.71$, $SD = 1.13$) rather than from the high *informativeness* condition ($M = 3.37$, $SD = 1.07$), $t(200) = 2.21$, $p = .03$. As with *process satisfaction*, the differences between values across *assortment size* were non-significant in the low *informativeness* ($F(9, 91) = 1.22$, $p = .30$), but significant in the high *informativeness* condition ($F(9, 91) = 2.05$, $p = .04$). A significant linear relationship between *choice overload* and *assortment size* is existent for low ($r = .19$, $p = .03$) and high *informativeness* ($r = .37$, $p < .001$), whereas the linearity in the high *informativeness* condition is considerably higher. For the low *informativeness* condition, the results of a linear regression are marginally significant and a model would only be able to explain 4 % of variance ($R^2 = .036$, $F(1, 99) = 3.66$, $p = .059$). However, a linear model for the high *informativeness* condition is significant and is able to explain 14 % of variance in *choice overload* ($R^2 = .135$, $F(1, 99) = 15.46$, $p < .001$). In this condition, an increase in *assortment size* with 5 options each estimates an increase in *choice overload* by .03 ($\beta = .03$, $t(99) = 3.92$, $p < .001$).

As seen in Figure 7, *choice overload* values are higher for all *assortment sizes* in the low *informativeness* compared to the high i*nformativeness* condition, except for 20 and 50 options, where they are nearly equal.

As both preconditions seem to be reliable predictors for *choice overload*, a multiple linear regression is performed for *choice overload* as the dependent variable, and *assortment size* and *informativeness* as the independent variables (for the results see Table 5). The model is significant and can explain 9 % of the variance. When 5 options are added to an assortment, the *choice overload* rating would increase by .02, whereas an addition of 4 attributes to the options leads to a decrease of *choice overload* by -.38.

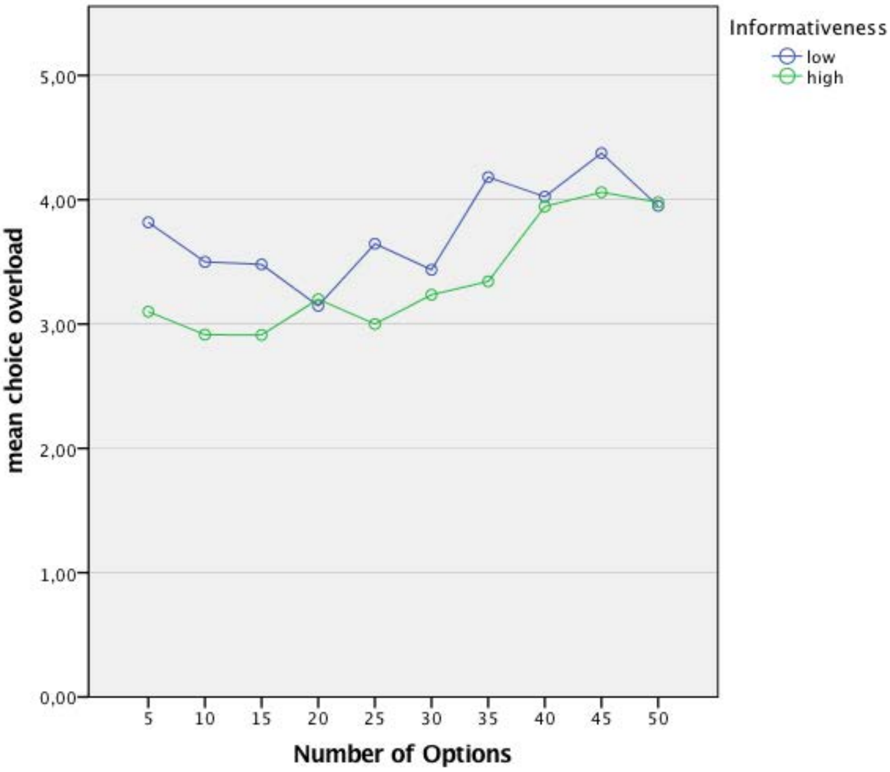


Figure 7. *Mean values for choice overload across assortment size for each informativeness condition.*

Table 5. *Results of a multiple linear regression, influence on choice overload*

variable	unstandardized	standardized	standard error
constant	3.501**		
assortment size	.022**	.273**	.006
informativeness	-.384*	-.173*	.150
R^2	.098		
corr. R^2	.089		
F(2,199)	10.83**		

*p < .05, **p < .001

3.2.3 The Influence of other Moderators on the Relationship between Assortment Size and the Choice-Making Process

Apart from the amount of presented attributes in the playlists, other moderators are suggested to have an influence on the relationship between assortment size and the choice-making process. As suggested in prior research (Schwartz et al., 2002), the *propensity to maximize* of participants is going to be analyzed. As the choice of a *dominant option* seems to have an impact as well (see section 3.1), a more detailed analysis on how it influences the relationship will follow too.

PROPENSITY TO MAXIMIZE. To see whether the propensity to maximize has an influence on the outcome variables of a choice-making process, an ANOVA is performed for each of the outcome variables with *propensity to maximize* as the factor. For *process satisfaction*, the differences between values is significant ($F(20, 181) = 1.85, p = .02$). No significant differences could be found for *negative effects* ($F(20, 181) = 1.17$, $p = .29$), *switching likelihood* ($F(20, 181) = 1.18$, $p = .28$) and *choice overload* ($F(9, 91) = 1.04$, $p = .42$) though. Looking at the correlations, only *process satisfaction* and *propensity to maximize* have a low significant linear relationship ($r = 0.13$, $p = .03$). Nevertheless, a linear regression reveals that propensity to maximize is no reliable predictor for *process satisfaction* levels ($R^2 = .017$, $F(1, 200) = 3.42$, $p = .066$). A look at the mean values of the outcome variables across *propensity to maximize* shows no plausible tendencies (Figure 8).

DOMINANT OPTION. As displayed in section 3.1, the choice of a known song resulted in lower levels of *choice overload*. Looking at the other outcome variables shows that levels of *process satisfaction* were also significantly higher when people chose a dominant option ($M = 4.69$, $SD = 1.19$) compared to those who did not ($M = 3.99$, $SD = 1.27$), $t(200) = 2.66$, $p = .008$. For *negative effects* and *switching likelihood* there are no significant differences between groups ($t(200) = -1.59$, $p = .11$; $t(200) = -1.18$, $p = .24$, respectively). When a *dominant option* was selected, there are no significant differences between values of each of the outcome variables across *assortment size*. As the choice of a dominant option could not be controlled in the experiment, the cases for this group are distributed very unequally. Looking at the frequencies shows that there are just a few people for every assortment size that chose a song they knew, for some sizes there are no people (see **Fehler! Verweisquelle konnte nicht gefunden werden.**). With an experimental group this small and unequal, no comparisons are possible between both conditions and no further analysis will take place.

Table 6. *Absolute frequencies of the selection of a dominant option across assortment size.*

assortment size	**5**	**10**	**15**	**20**	**25**	**30**	**35**	**40**	**45**	**50**	**total**
no	17	13	21	18	19	25	13	17	15	17	***175***
yes	0	0	5	5	3	3	5	2	3	1	***27***
total	**17**	**13**	**26**	**23**	**22**	**28**	**18**	**19**	**18**	**18**	***202***

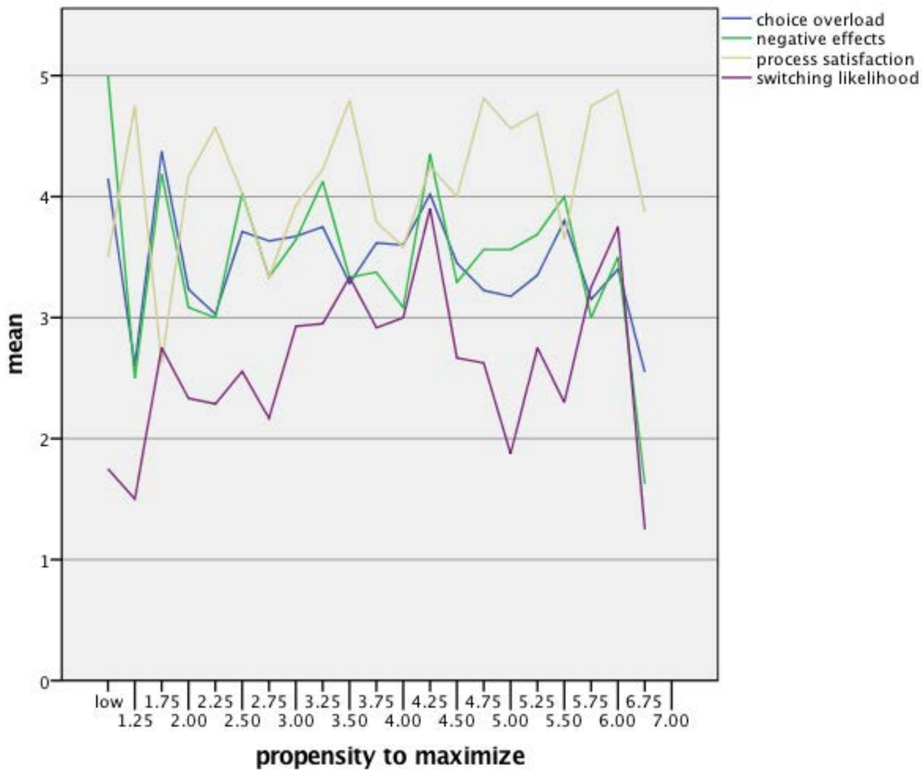


Figure 8. *Mean values for outcome variables across propensity to maximize.*

3.3 Music Listening Experience

This chapter deals with the outcome of the choice: the music listening experience. As seen in the sections above, *assortment size* and *informativeness* are good predictors for choice overload. However, the topic of interest is how the level of choice overload one experiences influences the actual outcome of the choice. It is therefore tested if the music listening experience is a function of choice overload. As all of the investigations on the *choice overload* measure revealed significant results, it is used as the independent variable for further examinations.

All *choice overload* values from 1.2 until 6.0 (with increments in .2) are represented at least twice, with middle values being more frequent than very low and high values. Ideally, all frequencies should be equal, but as *choice overload* could not be directly manipulated, it is enough to have each of the values represented at least twice.

The reliability of the scale *listening experience*, in which all the items intended to measure the music listening experience, is very high (Cronbach's α = .93, for further information see section 2.5.2). The correlation values in Table 7 show that they all have a really high, significant linear relationship, suggesting that they are all measuring the same construct. The measure *listening* *experience* seems to highly represent all of the tested scales and therefore it is used as the dependent variable in further examinations.

As *expectation disconfirmation* is another item to examine the outcome of the choice, it is looked at as well. In order to make the evaluation easier, the values have to be inverted because originally, a high value means low *expectation disconfirmation* ("much better than expected"). With the inversion, a low value means low *expectation disconfirmation* and vice versa. It could not be aggregated into *listening experience*, as it used a 9-point rating scale instead of a 7-point scale (as the other items). However, both of the outcome variables have a strong linear relationship ($r = -.82$, $p < .001$), suggesting that they measure the same construct (with *expectation disconfirmation* inversed).

***Table** 7. r-values of a Pearson correlation*

	listening experience	aesthetic properties	subjective value	outcome satisfaction
listening experience	1			
aesthetic properties	,938**	1		
subjective value	,922**	,767**	1	
outcome satisfaction	,930**	,862**	,767**	1

***p < .001*

Participants had to rate their music listening experience on a slider from 0 to 100 at the end of the outcome assessment. A correlation between this variable and *listening experience* shows that they are highly correlated (r = .85, $p < .001$), a finding that supports that all the items intended to measure the music listening experience, which were aggregated into *listening experience,* prove to be good items.

3.3.1 The Impact of Choice Overload on Music Listening Experience

LISTENING EXPERIENCE. The mean values of *listening experience* differ statistically significant across choice overload, suggesting that there is a relationship between these variables ($F(24, 177) = 2.26$, $p = .001$). A correlation supports this assumption, showing that there is a moderate linear relationship ($r = -.35, p < .001$). A linear regression reveals a significant model as well, which is able to explain 12 % of variance of *listening experience* through *choice overload* ($R^2 = .121$, $F(1, 200) = 27.61$, $p < .001$). The model predicts a decrease in *listening experience* with an increase in *choice overload* ($\beta = -.43$, $t(200) = -5.26$, $p < .001$).

In Figure 9, this trend is observable as well. The standard error is for some values relatively high, which is a consequence of the low number of cases in some *choice overload* conditions.

For the low and high *choice overload* levels, the standard error is higher than in the middle scores, as there are less cases for them.

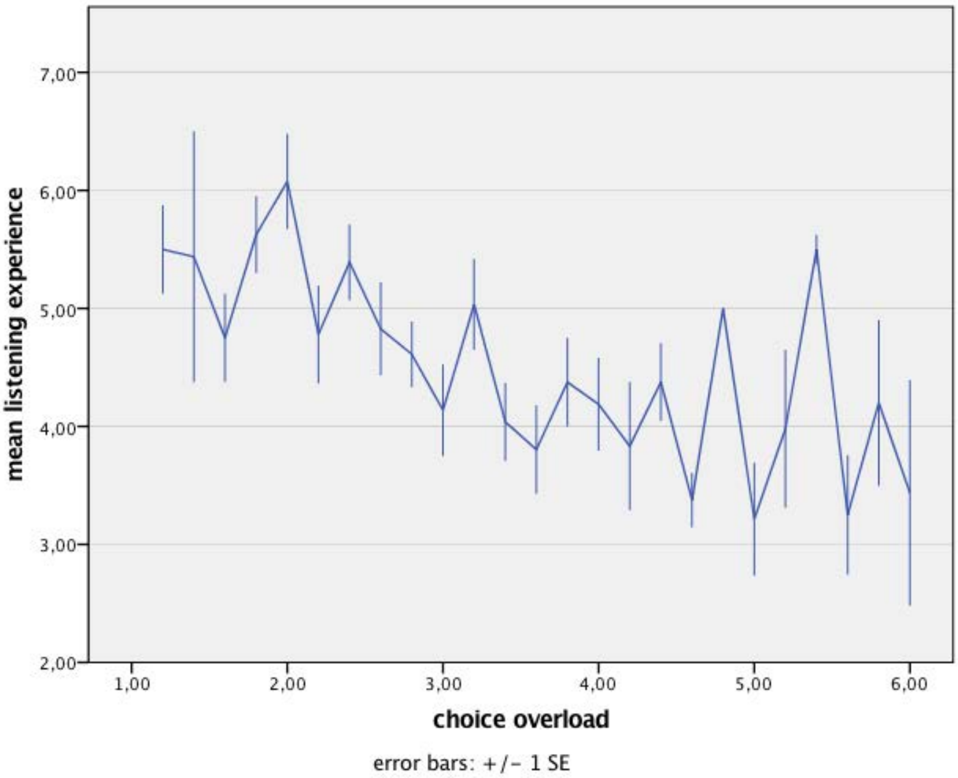


Figure 9. *Mean values for listening experience across choice overload.*

EXPECTATION DISCONFIRMATION. A significant effect of *choice overload* on *expectation disconfirmation* could not be found ($F(24, 177) = 1.52$, $p = .07$). However, a weak positive trend of *expectation disconfirmation* can be observed across *choice overload* ($r = .18$, $p = .009$). A linear model is significant, but could just explain 3 % of variance ($R^2 = .028$, $F(1, 200) = 5.70$, $p = .02$). The model predicts an increase in *expectation disconfirmation* with an increase in *choice overload* ($\beta = .33$, $t(200) = -2.39$, $p = .02$). Looking at the mean values and standard errors (Figure 12 in the appendix) across *choice overload* shows that the linear trend is not as clear as in *listening experience,* with standard errors being even larger.

3.3.2 The Influence of Informativeness on the Relationship between Choice Overload and Music Listening Experience

As *choice overload* levels were different between the two *informativeness* conditions, it is appropriate to further investigate on *listening experience* when the cases are divided in low and high *informativeness*. As all observations will be split into *informativeness* conditions, the cases in each *choice overload* condition will be less than before.

LISTENING EXPERIENCE. Participants facing playlists with low *informativeness* reported similar levels of *listening experience* ($M = 4.51$, $SD = 1.39$) as participants facing a playlist with high *informativeness* ($M = 4.26$, $SD = 1.34$), $t(200) = 1.33$, $p = .19$. *Listening experience* differed significantly across *choice overload* for low ($F(21, 79) = 1.81$, $p = .03$) and high ($F(24, 76) = 1.72$, $p = .04$) *informativeness*. Correlation values show a significant negative relationship for both conditions ($r = -.29, p = .002$, $r = -.46, p < .001$; low, high). A linear model with *choice overload* as the predictor for *listening experience* in the low *informativeness* condition is significant, but could only explain 8 % of variance ($R^2 = .082$, $F(1, 99) = 8.87$, $p = .004$). It predicts a moderate decrease in *listening experience* with an increase in *choice overload* ($\beta = -.35$, $t(99) = -2.98$, $p = .004$). The same relationship can be observed in the high *informativeness* condition, but with a higher ability to explain the variance of *listening experience* ($R^2 = .211$, $F(1, 99) = 26.46$, $p < .001$) and a stronger decrease in the dependent variable ($\beta = -.58$, $t(99) = -2.98$, $p < .001$). The standard errors are for both conditions relatively high (see Figure 13 in the appendix).

As both *choice overload* and *informativeness* seem to significantly influence *listening experience*, a model is required to predict *listening experience* levels. For that, a multiple linear regression is performed. Table 8 shows that a model which takes both predictors into account can explain 13.5 % of variance of *listening experience*. It also predicts a decrease in *listening experience* when *choice overload* and *informativeness* increase.

Table 8. *Results of a multiple linear regression, influence on listening experience*

variable	unstandardized	standardized	standard error
constant	6.625**		
choice overload	-.458**	-.372**	.082
informativeness	-.412*	-.151*	.181
R^2	.143		
corr. R^2	.135		
F(2,199)	16.67**		

**p < .05, **p < .001*

EXPECTATION DISCONFIRMATION. As in *listening experience*, there is no significant difference between low ($M = 4.18$, $SD = 2.20$) and high *informativeness* ($M = 4.58$, $SD = 2.21$) for *expectation disconfirmation*, $t(200) = -1.31$, $p = .19$. The differences in values across *choice overload* are non-significant in both groups ($F(21, 79) = 1.11$, $p = .35$, $F(24, 76) = 1.15$, $p = .31$; low, high). In the low *informativeness* condition, no significant linear relationship between *expectation disconfirmation* and *choice overload* could be found ($r = .13$, $p = .09$), whereas with high *informativeness,* such a relationship is existent ($r = .24$, $p = .008$). A linear model is significant and could explain 6 % of variance of *expectation disconfirmation* in the high *informativeness* condition ($R^2 = .057$, $F(1, 99) = 5.98$, $p = .016$). It predicts a positive linear relationship between *expectation disconfirmation* and *choice overload* ($\beta = .49$, $t(99) = 2.45$, $p = .016$). For low *informativeness*, such a model is non-significant.

3.3.3 The Influence of other Moderators on the Relationship between Choice Overload and Music Listening Experience

As in section 3.2.3, the influence of a *dominant option* will not be evaluated for the same reason (too few cases). However, even though there are no remarks in prior research, the influence of *propensity to maximize* on the relationship between *choice overload* and *listening experience* will be examined as the data allows it.

PROPENSITY TO MAXIMIZE. The results of an ANOVA show, that *propensity to maximize* most likely has an influence on *listening experience*, as the values differ significantly across moderator levels ($F(20, 181) = 2.79$, $p < .001$). This influence is supported by a significant correlation value ($r = .22$, $p = .001$). A linear regression reveals a significant model, with a low ability to explain variance of *listening experience* ($R^2 = .048$, $F(1, 200) = 10.11$, $p = .002$). The model predicts an increase in *listening experience* with rising levels of *propensity to maximize* ($\beta = .24$, $t(200) = 3.18$, $p = .002$).

As *propensity to maximize* seems to be a further predictor for *listening experience*, it is included to extend the model of the multiple linear regression. The results are presented in Table 9. With *propensity to maximize* included in the model, 16.5 % of variance of *listening experience* can be explained through the predictors. An increase in *choice overload* and *informativeness* mitigate *listening experience*, whereas increasing levels of *propensity to maximize* result in higher levels of *listening experience*.

Table 9. *Results of a multiple linear regression, influence on listening experience*

variable	unstandardized	standardized	standard error
constant	5.813**		
choice overload	-.435**	-.353**	.081
informativeness	-.404*	-.148*	.178
propensity to maximize	.199*	.185*	.070
R^2	.178		
corr. R^2	.165		
F(3,198)	14.24**		

**p < .05, **p < .001*

4 Discussion

The analyzed data provides good material for further interpretation and comparisons with the results of prior research.

4.1 Choice-Making Process

4.1.1 The Impact of Assortment Size on the Choice-making Process

The results show, that *process satisfaction* is highest at 25 and 30 options, and lowest at 40 options. No linear or curvilinear relationship could be determined, because after 40 options, *process satisfaction* increases again. The standard error is highest for the values at 45 and 50 options, suggesting that the results could turn out different in another sample and in the wider population. If the values in those conditions would be lower, an inverted U-shape would be present and the results of Reutskaja & Hogarth (2009) could be confirmed, suggesting that *process satisfaction* and *assortment size* have a curvilinear relationship. It is also inconceivable that *process satisfaction* will continue to increase as the number of options in an assortment increases. Nevertheless, *process satisfaction* shows a peak in middle sized choice sets. A finding that confirms a part of hypothesis 1. (see section 1.3), suggesting that there will be a peak of enjoyment and anticipated satisfaction which will not be at the smallest and largest assortment.

For *negative effects*, a linear relationship could be identified which estimates an increase in difficulty and frustration in the choice-making process with an increase in *assortment size*. A finding that supports hypothesis 1., as well as the general choice-overload hypothesis. However, the linear model is not very strong ($R^2 = .069$), suggesting that other factors need to be evaluated to explain a larger proportion of variance. Those factors could be the stimuli, presentation format or person-related factors. The linear model could be not that strong because of the minimum in *negative effects* at an assortment size of 15 options. At this point, the difference between *process satisfaction* and *negative effects* is the largest, suggesting that the choice-making process is most pleasant when presented 15 options. Between 30 and 35 options, the courses of these two variables cross, which means that *negative effects* outweigh *process satisfaction* (the difference between variables starts to get negative). This point can be regarded as the threshold for choice overload. As for smaller choice sets, the difference between these measures is quite small as well (5 and 10 options) until it reaches the maximum at

15 options, it indicates that there could be reverse effect of choice overload: not having too much to choose from but too few to choose from. A finding that questions prior research, which assumed that in very small choice sets no negative effects are present (for presented number of options in small vs. large assortments, see section 1.1.1). The difference between *negative effects* and *process satisfaction* is relatively small for 5 and 35 option assortments, suggesting that the choice-making process has similar effects on participants in both set sizes. This is a finding that could explain why Scheibehenne et al. (2009) were not able to detect a choice overload effect with 6 vs. 30 options in the music domain, as in middle sized assortments and larger assortments the strongest effects could be found. The results support the notion that more and larger assortment sizes need to be examined in further research regarding music steaming because small choice sets can have a negative effect on the choice-making process as well.

Switching likelihood and *assortment size* showed a linear relationship as well, but only a weak one with no significant differences between mean values. It demonstrates that frustration and difficulty are more sensitive measures to react to a change in *assortment size*, suggesting that they are better predictors for choice overload. *Switching likelihood* has been an outcome-based indicator in prior research, which could be answered on only two dimensions (Chernev, 2003) rather than on a rating scale. Through the different implementation in this experiment a strong effect could have been attenuated.

Taking all of the above measures together, a *choice overload* scale could be developed. *Process satisfaction* had to be inverted to create a meaningful construct. The relationship of *choice overload* and *assortment size* proved to be linear as well, with an increase in *assortment size* leading to higher levels of *choice overload.* However, the linear model is not very strong (R^2 = .068), once again suggesting that *assortment size* is not able explain enough variance alone and that other factors play a role as well. The course across mean values for each *assortment size* is similar to that of *negative effects*, with choice overload being higher at choice sets with 5 and 10 options and a minimum at 15 options, afterwards it increases. This supports the idea that the choice-making process is most pleasant at 15 options (a finding that is supported by the results of the master thesis by Miguel Reyes Botello (2021), who found that choice overload is the weakest at 16 options when selecting music from playlists). After 15 options, *choice overload* increases steadily with a really strong increase between 30 and 35 options (difference of neighboring mean values is the largest), suggesting that the choice overload effect gets stronger at this point and the negative effects outweigh. A peak is observable at 45 options, at 50 options it is smaller again. A finding which is most likely due to the sample, since it is inconceivable that from this point on the *choice overload* level continues to decrease. Without the decrease at 50 options, a curvilinear relationship in an U-shape of *choice overload* across *assortment size* could be present. Comparing the mean values leads to the discovery, that at 5 options the choice overload level is even higher than at 25 and 30 options. A finding that once again illustrates the need for more assortment sizes in future experiments. The largest difference of mean values is between 15 and 45 options, suggesting that 6 vs. 30 options (Scheibehenne et al., 2009) are too small to examine choice overload in the music domain.

The results show that for all *process satisfaction*, *negative effects* and hence *choice overload* a curvilinear relationship across *assortment size* is most likely present. This could also be the reason why the linear model is not that strong with only small coefficients. They also demonstrate that the consequences of large assortments can be also present in very small assortments (which indicates that there can be too few options to choose from). Participants had the most pleasant experience at 15 options, a further increase in assortment size after this point leads to higher choice overload levels. Choice overload most likely occurs between 30 and 35 options, as at this point the negative effects outweigh the positive effects. It could also be demonstrated that difficulty, frustration, enjoyment in the choice-making task and anticipated satisfaction prove to be good measures to examine choice overload (Iyengar & Lepper (2000) were also able to show significant results with these items).

4.1.2 The Influence of Informativeness on the Relationship between Assortment Size and the Choice-Making Process

In the section before, the level of *informativeness* of the playlists has not been considered in the evaluation. To see how it affects the relationship between assortment size and choice overload, the results of section 3.2.2 will be discussed in this part.

When playlists are presented with six attributes to each song, *process satisfaction* is significantly higher than when songs are just presented with title and artist. That does not support hypothesis 1., which suggests higher *choice overload* levels (hence lower *process satisfaction*) with an increase in number of attributes. Furthermore the findings of Greifeneder et al. (2010) could not be supported, showing lower satisfaction levels with a higher number of attributes. *Process satisfaction* levels of high and low *informativeness* meet at 40 options, with a further increase in *assortment size* they are very similar. However, it is higher with high *informativeness* for all assortment sizes before the 40 option mark. This finding reinforces the previous results that the presence of categories (here it could be the musical style) makes the choice-making process easier and more enjoyable (Huffman & Kahn, 1998; Mogilner et al., 2008). However, the results show that this is only until 40 options, suggesting that an increase in *assortment size* from that point with a high level of *informativeness* might bring too much information and therefore attenuates *process satisfaction*.

For *negative effects*, no significant differences between assortment size conditions could be found. In both conditions, the linearity between *assortment size* and *negative effects* is still existent. This indicates that a change in number of attributes has no influence on frustration and difficulty in the choice-making process, but on enjoyment and anticipated satisfaction. It could be that anticipated satisfaction is the crucial factor in this construct, as participants (in the *high informativeness* condition) were more informed about the presented songs and could tell more if there are going to be satisfied with their choice.

Comparing the mean values across *assortment size* of *process satisfaction* and *negative effects* in each *informativeness* condition, shows that with just two attributes, the *negative effects* outweigh *process satisfaction* at 5 options. This suggests that participants experience the same effects of choice overload, but with a very small assortment. The highest difference of mean values is at 20 options, suggesting that this *assortment size* was the most pleasant to choose from with low *informativeness*. In this condition, the courses of *process satisfaction* and *negative effects* cross between 30 and 35 options. After this point, negative effects are higher than process satisfaction. This suggests that for playlists with low *informativeness*, the threshold for choice overload lies between 30 and 35 options.

In the high *informativeness* condition, the difference between mean values of *process satisfaction* and *negative effects* is relatively high for all assortment sizes from 5 to 30 options, showing that the negative consequences of small playlists are not present in this condition (like in the other group – could be related to presence of categories again). The highest difference is at 10 and 25 options, suggesting that these assortment sizes were the most pleasant to choose from. From 25 options up, *process satisfaction* decreases and *negative effects* increases, until they meet between 35 and 40 options. This suggests that playlists with a higher number of attributes shift the threshold for choice overload to larger assortments.

Switching likelihood levels did not differ significantly between both *informativeness* conditions. This is not surprising, as *negative effects* showed no difference as well. It means that negative consequences do not change with an increase in number of attributes, but it makes the choice-making task more enjoyable and raises anticipated satisfaction until a certain point (40 options).

Analogous to the change in *process satisfaction*, *choice overload* levels were significantly higher in the low *informativeness* condition. This is not a surprising result, as *negative effects* and *switching likelihood* were similar across *informativeness* conditions but *process satisfaction* was lower. A linear relationship between *assortment size* and *choice overload* is in both conditions present. For high *informativeness,* the linear model is stronger ($R^2 = 0.14$) than that for low *informativeness* ($R^2 = 0.07$), which suggest that with a higher number of attributes, *assortment size* seems to be a better predictor for *choice overload* and there is not such a large need for other factors to explain the variance.

Overall, *choice overload* levels were higher in the low *informativeness* condition, with a minimum at 20 option. At this point, both *choice overload* courses meet. This suggests that when presented 20 options, there is no difference if songs are described by two or six attributes. When two attributes were presented, the difference between neighboring mean values was highest between 30 and 35 options (as in the case when *informativeness* was not considered), indicating that this is the point when choice overload occurs. Both curves meet at 40 options again, suggesting that from this assortment size upwards, there is no difference in *choice overload* levels for low and high *informativeness*. This finding supports the idea that after a certain set size is reached, the information is too much to process to sustain a high level of *process satisfaction*.

With high *informativeness*, the lowest values for *choice overload* are at 10 and 15 options, after this an increase in *assortment size* leads to an increase in *choice overload*. It demonstrates that the most pleasant set sizes to choose from have 10 or 15 options in them. The highest difference between neighboring mean values is between 35 and 40 options, suggesting that this is the point where choice overload occurs. As mentioned before, from 40 options upwards there is no difference between *choice overload* levels, meaning *informativeness* has no influence anymore from this point. It gives the idea that participants who feel not informed (through a low number of attributes) and the amount of information (through a high number of attributes) outweigh each other at this point. It would be interesting to examine larger choice sets in this regard, because it is imaginable that *choice overload* levels would be higher in the high *informativeness* condition compared to the low *informativeness* condition for really large sets, as the amount of information increases rapidly with an increase in *assortment size*.

The existence of minimums for both conditions in middle sized assortments suggests that for both *informativeness* levels, a curvilinear relationship between *assortment size* and *choice overload* might be present.

The model of the performed multiple linear regression predicts a decrease of *choice overload* with an increase in number of attributes. However, this finding is questionable as at a certain point it would be too much information to process and thus exceed the cognitive resources, which basically is the definition of choice overload. The difference between two and six attributes is relatively large and it would be interesting to see how the behavior of the participants changes when three or four attributes are presented to each song. It could be that process satisfaction could even be higher when just the right (amount of) attributes are presented. This idea is analogous to the findings of Keller & Staelin (1987), who demonstrated an inverted U-shape between number of attributes and choice accuracy. It would also be interesting to examine which of the attributes (musical style or GMBI adjectives) actually mitigated the choice overload effect.

The results of this section suggest that in the music domain, choice overload can be attenuated by adding more attributes to the songs of a playlist. Nevertheless, this only works until a certain assortment size is reached, as afterwards there is no difference in in *choice overload* levels between both conditions.

4.1.3 The Influence of other Moderators on the Relationship between Assortment Size and the Choice-Making Process

As shown in chapter 1.1.2, there are a lot more conditions which attenuate or amplify choice overload as a function of assortment size. In the experiment, the *propensity to maximize* was recorded for each participant (as a person-related factor). Also the choice of a *dominant option* was recorded.

No effect of *propensity to maximize* on the dependent variables could be detected. A reason for this could be the low reliability of the scale ($\alpha = .63$), which makes it hard to come to meaningful conclusions. A possible problem could be that just five items from the original maximizing scale by Schwartz et al. (2002) were included. However, prior studies which included the whole scale reported similar issues (Scheibehenne et al., 2009). Hypothesis 3. can therefore neither be confirmed nor refuted.

The choice of a dominant option surely has an influence, as significant differences between groups were present in *choice overload* and *process satisfaction*. People who chose a dominant option reported higher levels of *process satisfaction* and lower levels of *choice overload*. This is a finding that supports prior research, suggesting that the availability of a dominant option from a large assortment decreases the cognitive effort and increases the likelihood of purchasing from a large assortment, which is an indicator for low choice overload (Chernev, 2006; Oppewal & Koelemeijer, 2005). The variable could not be controlled in the experiment, because even though rather unfamiliar songs were chosen as stimuli, some of the participants knew songs before and chose them. As this group is relatively small compared to the group which did not choose a *dominant option*, it is hard or even not possible to compare the outcome of both groups. Hypothesis 5. can be partly confirmed, as the choice of a dominant option really seems to lower choice overload levels. This assumption needs further investigation though.

Future experiments examining choice overload in the music domain should include this moderator, as it surely has an influence on the choice-making process.

4.2 Music Listening Experience

As the title of this work suggests, the core of this experiment is to evaluate the influence of choice overload on the music listening experience. Therefore, the dependent variable of the first part of the evaluation has been converted to the independent variable of the second part of the evaluation. This brings some disadvantages, the biggest of them being the frequencies in each condition. They are not evenly distributed and some of them are really low ($N = 2$). Because of that, it is expected that the results are not transferrable to the wider population and that the same experiment with a different sample will get other results. This assumption is also confirmed by the relatively large standard errors, especially for very high and low levels of *choice overload*. Nevertheless, it gives a first hint on how choice overload influences music listening experience, at least for this sample.

4.2.1 The Impact of Choice Overload on Music Listening Experience

A significant relationship between *choice overload* and *listening experience* could be revealed. It is negative and linear and a model is able to explain 12 % of variance of *listening experience* through *choice overload*. It predicts a decrease in *listening experience* with an increase in *choice overload*. This is a finding that is supported by the choice overload hypothesis and prior research, which suggest a lower outcome satisfaction when choice overload is present (Botti & Iyengar, 2004; Iyengar & Lepper, 2000). The results demonstrate that not only outcome satisfaction is influenced by *choice overload*, but also how the music is perceived (aesthetic properties and subjective value). This suggests that the occurrence of choice overload has an impact on how the product is perceived. Hypothesis 2. can be confirmed for this sample.

The results also show that *choice overload* has an impact on *expectation disconfirmation*, with an increase in *choice overload* resulting in higher levels of *expectation disconfirmation*. It could be that the people who experience choice overload have higher expectations of a product, which are not met. Maybe this is because people expect the outcome of an unpleasant decision task to be really good. The expectation is with a high level of choice overload probably so high, that one gets disappointed, regardless of the products performance. This then leads to a lower listening experience (which explains the strong correlation between *listening experience* and *expectation disconfirmation*). The results are supported by the theory of Oliver (2003), who describes satisfaction as the relation between expectation and a product's performance. Large assortments (and probably also choice overload) rise expectations but the performance of the product stays the same, which means that satisfaction with a product is dependent from the size of the assortment it was chosen from. This idea explains why choice overload has an influence on how the music is perceived, regardless of the actual music.

4.2.2 The Influence of Informativeness on the Relationship between Choice Overload and Music Listening Experience

As *informativeness* has an influence on *choice overload*, it is expected to have an influence on *listening experience* as well. It is expected that *listening experience* will decrease with an increase in *choice overload* in both *informativeness* conditions, but that the decrease will be less steep in the high *informativeness* condition.

There is no significant difference between low and high *informativeness* in *listening experience* levels. However, for each condition, a significant linear model could be calculated. It is not very strong for low *informativeness* (R^2 = .08), but for high *informativeness* (R^2 = .21). This suggests that in the low *informativeness* condition, more factors play a role on the relationship between *choice overload* and *listening experience*. By looking at the regression coefficient, one can see that the slope in the low *informativeness* condition is smaller (β = -.35) than in the high *informativeness* condition (β = -.58). The findings suggest that an increase in *choice overload* in the low *informativeness* condition is less drastic rather than in the high *informativeness* condition. This finding is surprising, as the addition of attributes resulted in lower *choice overload* levels. One might think that it would thus lead to a smaller decrease in *listening experience*. It could be that when more attributes are presented, the expectation of a song gets higher which mitigates the listening experience. An effect that is not present when just the title and artist are presented. In this case participants are probably more open about the music and are satisfied way faster. These findings show that the traits of a choice set can influence the outcome with a choice in other ways than it influences the choice-making process. A crucial finding to support the idea that the choice-making process and the outcome with the choice should be evaluated separately. A linear model which takes *choice overload* and *informativeness* for listening *experience* as predictors into account, estimates a decrease in *listening experience* levels with an increase in *choice overload* and *informativeness* (see Table 8). This is different in for the choice-making process, where an increase in *informativeness* leads to a decrease in *choice overload* (which makes the choice-making process more pleasant).

The results do not confirm the hypothesis 2.a), as the *listening experience* levels do not differ between *informativeness* conditions. Additionally, the slope in the low *informativeness* condition is less steep instead of that in the high *informativeness* condition.

Expectation disconfirmation does not differ as well across informativeness conditions. A linear relationship could only be detected in the high *informativeness* condition. It predicts an increase in *expectation disconfirmation* with an increase in *choice overload*. This supports the idea that expectations rise with a higher level of *choice overload*. However, a statement regarding differences between both *informativeness* conditions is not possible, as there are no meaningful results in the low i*nformativeness* condition.

It should be mentioned, that for each condition, the frequencies are even lower than in the examination before (no consideration of *informativeness*), as the cases were split into the two *informativeness* conditions. This makes the results even less reliable.

4.2.3 The Influence of other Moderators on the Relationship between Choice Overload and Music Listening Experience

As the data allows it, the influence of *propensity to maximize* was analyzed. Contrary to the examined influence on *choice overload* (which could not be revealed), a significant relationship could be detected for *listening experience*. A linear model predicts an increase in *listening experience* levels with an increase in *propensity to maximize*. It also seems to be a significant predictor of *listening experience*, next to *choice overload* and *informativeness*, in a linear model. This finding refutes hypothesis 3, which expected a decrease in *listening experience* with an increase in *propensity to maximize*. How is this possible? First, this relationship is present regardless of the level of *choice overload*. It could be that people who tend to maximize make their choice very thoughtful, and go through all the options to be really sure to get the best option. This might make the choice-making process more unpleasant (no evidence found!) but the outcome with the choice even better. On the other hand, this procedure would raise expectations of the individual, which would lead to a lower listening satisfaction. This is definitely an issue that could be addressed in further research.

4.3 Why Choice Overload as Independent Variable?

An appropriate question is why *choice overload* was used as the independent variable and not *assortment size*, as *choice overload* seems to be a function of *assortment size*?

First of all, there are no significant relationships between *assortment size* and the music listening experience scales, which makes the creation of a meaningful construct not possible. But with *choice overload*, all of the outcome variables have moderate linear relationships (see table Table 10). Choice Overload is a construct that is an interplay of many factors, some of them still unknown to this point. Assortment size is just a number and is probably not a good predictor for the perception of music. The number of options in an assortment and the number of attributes are not enough to make a statement about music listening experience, as the step in the middle – the choice-making process – would not be considered. And this part has probably the biggest influence on the outcome with the choice. A notion that supports the idea of a separate evaluation of the choice-making process and the outcome with the choice (Scheibehenne et al., 2010). So, in the second part of the evaluation, the influence of the choice-making process on the listening experience was examined. For the sample, *choice overload* seems to be a good predictor for *listening experience*, which assortment size could not be. This shows that the experience with a product is not just a function of different sized assortments, but of a more complex construct.

Table 10. *r-values of a Pearson correlation*

	listening experience	aesthetic properties	subjective value	outcome satisfaction
choice overload	-.35**	-.32**	-.33**	-.32**
assortment size	-.07	-.11	-.03	-.05

**$p < .01$

As choice overload is not directly manipulable, the results are not really transferrable to a real life situation when creating a playlist. However, if it is known how the music listening experience is influenced by the choice-making process, one could (with enough knowledge) create a playlist with factors to control for choice overload and thus influence the listening experience.

5 Conclusion

The purpose of this study was to examine the influence of different sized assortments on the choice-making process of consumers in the music domain. Furthermore, it was of interest to see how the choice-making process affects the experience with the selected music. This interplay seems especially relevant when looking at the enormous amount of music which is available online through numerous music streaming services.

As prior research found out, such a vast catalogue of options can be detrimental to consumers, lowering satisfactions levels and increasing the difficulty of a choice task. These findings challenge the conventional knowledge that “more is better”. Iyengar & Lepper (2000) were the first to execute an extensive experiment dealing with the negative consequences of large assortments. They showed that when choosing from large assortments, participants experienced higher levels of regret and lower satisfaction with the chosen option and found the decision process to be more difficult, but also enjoyable. Since this publication depicts the foundation of choice overload research, a large number of publications on this phenomenon followed. The situation in which a decision task exceeds the cognitive resources of an individual, is called choice overload.

Taking knowledge from prior research, an online listening experiment was developed to learn more about the choice overload phenomenon in the music streaming context.

A problem which is prominent in older experiments is that just two assortment sizes were looked at. This limits the results, as it cannot be examined what happens between these sizes. This is a problem that has been addressed in this work, as it takes ten assortment sizes into account. Furthermore, possible moderators of choice overload were discussed in terms of transferability to music streaming. A change of number in attributes was added to the experimental conditions. This led to a 10x2 experimental design. In order to assess the choice-making process and the outcome with the choice (music listening experience) separately, the items to evaluate the dependent variables were taken from different publications.

Individuals who participated first got assigned to one of the experimental conditions. They saw a playlist (with a specific size and number of attributes) and had to choose one song from it. Subsequently, they had to rate the choice-making process in terms of process satisfaction, negative effects and switching likelihood. After this, they had to listen to a two-minute sample of the song they selected and rate their listening experience regarding aesthetic properties and subjective value of the music, outcome satisfaction and expectation disconfirmation. This separate evaluation of choice-making process and outcome with the choice was implemented as participants were not influenced by the outcome of their choice when evaluating their choice-making process. It also allowed to see whether the experience with the decision task has and influence on the listening experience. After evaluating the first choice, participants got assigned to a new experimental condition and had to run through the process again with a different sized playlist with another level of informativeness.

To rule out possible moderating effects, participants were asked to indicate the perceived attractiveness of the choice options and whether they chose a song they already knew. As option attractiveness and the existence of a dominant option have an impact on the relationship between assortment size and choice overload.

The results show that process satisfaction and assortment size most likely have a curvilinear relationship, which supports the results of Reutskaja & Hogarth (2009). The maximum of process satisfaction is at 25 and 30 options. Negative effects also have a curvilinear relationship with a minimum at 15 options. At this point the difference between process satisfaction and negative effects is the largest, suggesting that this is the ideal playlist size. The point when choice overload occurs can be regarded as the point where negative effects outweigh process satisfaction, which is between 30 and 35 options. The course of an aggregated choice overload score support these findings. The results further suggest that the effects of choice overload can also be present in relatively small options, as the choice overload score is higher in these dimensions as in middle sized dimensions. Looking at the difference between 5 and 30 options (which are common set sizes in prior research) shows that there is no difference in choice overload levels. This finding makes it clear, why Scheibehenne et al. (2009) were not able to find a difference between these sizes. After 30 options, choice overload increases rapidly. The results demonstrate the need for larger and more assortments to be looked at in future experiments.

When divided in low and high informativeness, differences between both conditions are present. Process satisfaction is higher with more attributes, which supports the knowledge that the presence of categories makes the decision task more enjoyable (Huffman & Kahn, 1998; Mogilner et al., 2008). In negative effects there is no difference between informativeness conditions, suggesting that a change in number of attributes has no influence on frustration and decision difficulty. With six attributes, the most pleasant choice experience was at 10 and 20 options, and the point where choice overload kicks in between 35 and 40 options (a shift to larger assortments). When just title and artist are presented, the best choice-making experience is at 20 options, the choice overload threshold between 30 and 35 options. From 40 options upwards, there is no difference between high and low informativeness, suggesting that the information is too much to process from this point on. The choice overload levels were also significantly higher in the high informativeness condition. This finding refutes the results of Greifeneder et al. (2010), who found that with an increase in number of attributes the choice task gets more complex and unpleasant. A curvilinear relationship is present for both informativeness conditions, as choice overload has a minimum each in middle sized assortments.

The choice of a dominant option seemed to be a moderator but could not be controlled, therefore it should be looked at in future experiments.

In the second part of the evaluation, the influence of choice overload on music listening experience was assessed. The very low number of cases for each condition make the results very unreliable. It could be demonstrated, that music listening experience decreases with an increase in choice overload. Which shows that it not only has an influence on outcome satisfaction, but also on how the music is perceived. The presence of more attributes showed no significant differences, except that music listening decreases slower across assortment size in the low informativeness condition. This finding is reverse to the findings of the first part of the evaluation, where high informativeness increased the pleasantness of the choice-making process. This again shows the need for a separate evaluation. An experiment with a larger sample would be interesting, to see if the results are transferrable.

Overall, the results are highly interesting, especially from the first part of the evaluation. They show that there are specific playlist sizes which make the choice-making process most pleasant and that choice overload occurs in relatively large sets. The addition of attributes (musical style and adjectives) seems to make the choice-making process more pleasant and reduces choice overload. However, in playlists which include more than 40 options, there is no difference anymore. It implies that in playlists of music streaming services, the addition of attributes could mitigate the negative consequences of middle sized to large assortments, but not in very large assortments.

The second part of the evaluation shows that an increase in choice overload attenuates music listening experience, which means that playlists have to be created in a way that choice overload is suppressed as much as possible to guarantee the best listening experience for the consumers. However, this is very difficult to realize, as all the factors and conditions to choice overload are still not explored and person-related factors also play a large role.

In the experimental design, it would have been helpful to include an item of the perceived control of participants over the music choice, as the results would have been comparable to the knowledge from the music psychology domain listed in chapter 1.2. In accordance to this, the emotions evoked by the music could be of interest as well.

For future experiments, it would also be good to include more informativeness conditions, as the gap between two and six attributes is quite large. It would be also interesting to see which attributes attenuate choice overload in the music domain.

The results are probably specific to the music streaming domain, as when buying physical audio carriers, the preconditions would be totally different. However, the experimental design is not very realistic, as when facing playlists from music streaming providers, consumers are normally able to have a quick listen to a song to decide whether they want to hear the whole track or not. They do not have to go through all the options visually, as they had to in the experiment. Regarding this, an interesting experimental design would come up, which would be more realistic. Furthermore, the influence of the song itself might play a role in the assessment of the listening experience. As most of the participants did not know the songs before, they did not know what they were choosing. Maybe they selected a song which they really did not like, even though they had a pleasant choice-making process. This would be crucial for the second part of the evaluation, but not for the first part. It shows that it makes sense to evaluate the choice-making process and the outcome with the choice separately.

This work is just a starting point for examining user experience in terms of choice processes in the music streaming context. As mentioned before, more conditions play a role in experiencing choice overload (which leads to a decrease in listening experience), which have to be examined in future experiments. This brings knowledge of how to create playlists which bring the best listening experience. We now know, that when choosing from unknown music, that there are certain playlist sizes which minimize choice overload. However, this might be different when music is known to the users or if they choose from recommended playlists constituted of highly attractive options. As selecting only one song from one playlist is not very realistic, it would be also interesting to see how the choice-making process turns out when users have to choose among assortments (playlists).

Literature

Anglada-Tort, M., Steffens, J., & Müllensiefen, D. (2018). Names and Titles Matter: The Impact of Linguistic Fluency and the Affect Heuristic on Aesthetic and Value Judgements of Music. *Psychology of Aesthetics, Creativity, and the Arts*. https://doi.org/10.1037/aca0000172

Arnold, S. J., Oum, T. H., & Tigert, D. J. (1983). Determinant Attributes in Retail Patronage: Seasonal, Temporal, Regional, and International Comparisons. *Journal of Marketing Research*, *20*(2), 149–157. https://doi.org/10.1177/002224378302000205

Babin, B. J., Darden, W. R., & Griffin, M. (1994). Utilitarian shopping value. *Journal of Consumer Research*, *20*(4), 644–657.

Baumol, W. J., & Ide, E. A. (1956). Variety in Retailing. *Management Science*, *3*(1), 93–101. https://doi.org/10.1287/mnsc.3.1.93

Berger, J., Draganska, M., & Simonson, I. (2007). The influence of product variety on brand perception and choice. *Marketing Science*, *26*(4), 460–472. https://doi.org/10.1287/mksc.1060.0253

Berlyne, D. E. (1960). *Conflict, arousal, and curiosity.* McGraw-Hill. https://doi.org/10.1037/11164-000

Bettman, J. R., Luce, M. F., & Payne, J. W. (1998). Constructive consumer choice processes. *Journal of Consumer Research*. https://doi.org/10.1086/209535

Blood, A. J., & Zatorre, R. J. (2001). Intensively pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences*, *98*(20), 11818–11823.

Boatwright, P., & Nunes, J. C. (2001). Reducing Assortment: An Attribute-Based Approach. *Journal of Marketing*, *65*(3), 50–63.

Bollen, D., Knijnenburg, B. P., & Graus, M. (2010). Understanding Choice Overload in Recommender Systems Categories and Subject Descriptors. *Matrix*, *121*(January 1994), 63–70. https://doi.org/10.1145/1864708.1864724

Botti, S., & Iyengar, S. S. (2004). The psychological pleasure and pain of choosing: When people prefer choosing at the cost of subsequent outcome satisfaction. *Journal of Personality and Social Psychology*, *87*(3), 312–326. https://doi.org/10.1037/0022-3514.87.3.312

Broniarczyk, S. M. (2008). Product Assortment. In C. P. Haugtvedt, P. M. Herr, & F. R. Kardes (Eds.), *Handbook of Consumer Psychology* (pp. 755–779). Lawrence Erlbaum Associates. https://doi.org/10.4324/9780203809570.ch30

Broniarczyk, S. M., Hoyer, W. D., & McAlister, L. (1998). Consumers' perceptions of the assortment offered in a grocery category: The impact of item reduction. *Journal of Marketing Research*, *35*(2), 166–176. https://doi.org/10.2307/3151845

Chernev, A. (2003). When More Is Less and Less Is More: The Role of Ideal Point Availability and Assortment in Consumer Choice. *Journal of Consumer Research*, *30*(2), 170–183. https://doi.org/10.1086/376808

Chernev, A. (2005). Feature complementarity and assortment in choice. *Journal of Consumer Research*, *31*(4), 748–759. https://doi.org/10.1086/426608

Chernev, A. (2006). Decision focus and consumer choice among assortments. *Journal of Consumer Research*, *33*(1), 50–59.

Chernev, A., Böckenholt, U., & Goodman, J. (2015). Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology*, *25*(2), 333–358. https://doi.org/10.1016/j.jcps.2014.08.002

Chernev, A., & Hamilton, R. (2009). Assortment size and option attractiveness in consumer choice among retailers. *Journal of Marketing Research*, *46*(3), 410–420. https://doi.org/10.1509/jmkr.46.3.410

Dhar, R. (1997). Consumer preference for a no-choice option. *Journal of Consumer Research*, *24*(2), 215–231. https://doi.org/10.1086/209506
Dhar, R., & Nowlis, S. M. (1999). The effect of time pressure on consumer choice deferral. *Journal of Consumer Research*, *25*(4), 369–384. https://doi.org/10.1086/209545
Diehl, K. (2005). When two rights make a wrong: Searching too much in ordered environments. *Journal of Marketing Research*, *42*(3), 313–322. https://doi.org/10.1509/jmkr.2005.42.3.313
Diehl, K., Kornish, L. J., & Lynch, J. G. (2003). Smart Agents: When Lower Search Costs for Quality Information Increase Price Sensitivity. *Journal of Consumer Research*, *30*(1), 56–71. https://doi.org/10.2139/ssrn.340040
Diehl, K., & Poynor, C. (2010). Great expectations?! Assortment size, expectations, and satisfaction. *Journal of Marketing Research*, *47*(2), 312–322. https://doi.org/10.1509/jmkr.47.2.312
Fasolo, B., Hertwig, R., Huber, M., & Ludwig, M. (2009). Size, entropy, and density: What is the difference that makes the difference between small and large real-world assortments? *Psychology and Marketing*, *26*(3), 254–279. https://doi.org/10.1002/mar.20272
Field, A. P. (2013). *Discovering Statistics Using IBM SPSS Statistics* (4th ed.). Sage Publications.
Goodman, J. K., & Malkoc, S. A. (2012). Choosing here and now versus there and later: The moderating role of psychological distance on assortment size preferences. *Journal of Consumer Research*, *39*(4), 751–768. https://doi.org/10.1086/665047
Gourville, J. T., & Soman, D. (2005). Overchoice and assortment type: When and why variety backfires. *Marketing Science*, *24*(3), 382–395. https://doi.org/10.1287/mksc.1040.0109
Greasley, A. E., & Lamont, A. (2011). Exploring engagement with music in everyday life using experience sampling methodology. *Musicae Scientiae*, *15*(1), 45–71. https://doi.org/10.1177/1029864910393417
Greenleaf, E. A., & Lehmann, D. R. (1995). Reasons for substantial delay in consumer decision making. *Journal of Consumer Research*, *22*(September), 186–199.
Greifeneder, R., Scheibehenne, B., & Kleber, N. (2010). Less may be more when choosing is difficult: Choice complexity and too much choice. *Acta Psychologica*, *133*(1), 45–50. https://doi.org/10.1016/j.actpsy.2009.08.005
Griffin, J. G., & Broniarczyk, S. M. (2010). The slippery slope: The impact of feature alignability on search and satisfaction. *Journal of Marketing Research*, *47*(2), 323–334. https://doi.org/10.1509/jmkr.47.2.323
Hadar, L., & Sood, S. (2014). When Knowledge Is Demotivating: Subjective Knowledge and Choice Overload. *Psychological Science*, *25*(9), 1739–1747. https://doi.org/10.1177/0956797614539165
Haynes, G. A. (2009). Testing the boundaries of the choice overload phenomenon: The effect of number of options and time pressure on decision difficulty and satisfaction. *Psychology and Marketing*, *26*(3), 204–212. https://doi.org/10.1002/mar.20269
Hirschman, E. C., & Holbrook, M. B. (1982). Hedonic Consumption: Emerging Concepts, Methods and Propositions. *Journal of Marketing*, *46*(3), 92–101.
Hoch, S. J., Bradlow, E. T., & Wansink, B. (1999). The variety of an assortment. *Marketing Science*, *18*(4), 527–546. https://doi.org/10.1287/mksc.18.4.527
Hotelling, H. (1929). Stability in Competition. *The Economic Journal*, *39*(153), 41–57. https://doi.org/10.2307/2224214
Huffman, C., & Kahn, B. E. (1998). Variety for sale: Mass customization or mass confusion? *Journal of Retailing*, *74*(4), 491–513. https://doi.org/10.1016/S0022-4359(99)80105-5
Inbar, Y., Botti, S., & Hanko, K. (2011). Decision speed and choice regret: When haste feels like waste. *Journal of Experimental Social Psychology*, *47*(3), 533–540. https://doi.org/10.1016/j.jesp.2011.01.011

Inman, J. J. (2001). The role of sensory-specific satiety in attribute-level variety seeking. *Journal of Consumer Research*, *28*(1), 105–120. https://doi.org/10.1086/321950
Iyengar, S. S., Jiang, W., & Huberman, G. (2004). How much Choice is too much: Determinants of Individual Contribution in 401k Retirement Plans. In O. S. Mitchell & S. Utkus (Eds.), *Pension Design and Structure: New Lessons from Behavioral Finance* (pp. 83–97). Oxford University Press. https://doi.org/10.1093/0199273391.001.0001
Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, *79*(6), 995–1006. https://doi.org/10.1037/0022-3514.79.6.995
Jacoby, J., Speller, D. E., & Kohn, C. A. (1974). Brand Choice Behavior as a Function of Information Load. *Journal of Marketing Research*, *11*(1), 63–69. https://doi.org/10.2307/3150994
Kahn, B., & Lehmann, D. (1991). Modeling Choice Among Assortments. *Journal of Retailing*, *67*(3), 274–299.
Kahn, B., Moore, W. L., & Glazer, R. (1987). Experiments in Constrained Choice. *Journal of Consumer Research*, *14*(1), 96–113. https://doi.org/10.1086/209096
Kahn, B., & Wansink, B. (2004). The Influence of Assortment Structure on Perceived Variety and Consumption Quantities. *Journal of Consumer Research*, *30*(4), 519–533. https://doi.org/10.1086/380286
Kahneman, D., & Snell, J. (1992). Predicting a changing taste: Do people know what they will like? *Journal of Behavioral Decision Making*, *5*(3), 187–200. https://doi.org/10.1002/bdm.3960050304
Karni, E., & Schwartz, A. (1977). Search theory: The case of search with uncertain recall. *Journal of Economic Theory*, *16*(1), 38–52. https://doi.org/10.1016/0022-0531(77)90122-3
Keller, K. L., & Staelin, R. (1987). Effects of Quality and Quantity of Information on Decision Effectiveness. *Journal of Consumer Research*, *14*(2), 200–213. https://doi.org/10.1086/209106
Krause, A., North, A., & Hewitt, L. (2014). Music Selection Behaviors in Everyday Listening. *Journal of Broadcasting and Electronic Media*, *58*(2), 306–323. https://doi.org/10.1080/08838151.2014.906437
Kreps, D. M. (1979). A Representation Theorem for Preference for Flexibility. *Econometrica*, *47*(3), 565–577.
Langer, E. J., & Rodin, J. (1976). The effects of choice and enhanced personal responsibility for the aged: A field experiment in an institutional setting. *Journal of Personality and Social Psychology*, *34*(2), 191–198. https://doi.org/10.1037/0022-3514.34.2.191
Lepa, S., Herzog, M., Steffens, J., Schoenrock, A., & Egermann, H. (2020). A computational model for predicting perceived musical expression in branding scenarios. *Journal of New Music Research*, *49*(4), 387–402. https://doi.org/10.1080/09298215.2020.1778041
Levav, J., & Zhu, R. (2009). Seeking freedom through variety. *Journal of Consumer Research*, *36*(4), 600–610. https://doi.org/10.1086/599556
Levy, M., & Weitz, B. A. (2006). *Retailing Management* (8th ed.). McGraw-Hill/Irwin.
Liljeström, S., Juslin, P. N., & Västfjäll, D. (2013). Experimental evidence of the roles of music choice, social context, and listener personality in emotional reactions to music. *Psychology of Music*, *41*(5), 579–599. https://doi.org/10.1177/0305735612440615
Malhotra, N. K. (1982). Information Load and Consumer Decision Making. *Journal of Consumer Research*, *8*(3), 419–430. https://doi.org/10.1086/208882
Markman, A. B., & Medin, D. L. (1995). Similarity and alignment in choice. *Organizational Behavior and Human Decision Processes*, *63*(2), 117–130. https://doi.org/10.1006/obhd.1995.1067
Messinger, P. R., & Narasimhan, C. (1997). A model of retail formats based on consumers’

economizing on shopping time. *Marketing Science*, *16*(1), 1–23. https://doi.org/10.1287/mksc.16.1.1
Mitchell, L. A., MacDonald, R. A. R., & Knussen, C. (2008). An Investigation of the Effects of Music and Art on Pain Perception. *Psychology of Aesthetics, Creativity, and the Arts*, *2*(3), 162–170. https://doi.org/10.1037/1931-3896.2.3.162
Mogilner, C., Rudnick, T., & Iyengar, S. S. (2008). The mere categorization effect: How the presence of categories increases choosers' perceptions of assortment variety and outcome satisfaction. *Journal of Consumer Research*, *35*(2), 202–215. https://doi.org/10.1086/588698
Morrin, M., Broniarczyk, S. M., & Jeffrey Inman, J. (2012). Plan format and participation in 401 (k) plans: The moderating role of investor knowledge. *Journal of Public Policy and Marketing*, *31*(2), 254–268. https://doi.org/10.1509/jppm.10.122
Oliver, R. L. (2003). Cognitive, Affective, and Attribute Bases of the Satisfaction Response. *Journal of Consumer Research*, *20*(12), 418–430.
Oppewal, H., & Koelemeijer, K. (2005). More choice is better: Effects of assortment size and composition on assortment evaluation. *International Journal of Research in Marketing*, *22*(1), 45–60. https://doi.org/10.1016/j.ijresmar.2004.03.002
Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance*, *16*(2), 366–387. https://doi.org/10.1016/0030-5073(76)90022-2
Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The Adaptive Decision Maker*. Cambridge University Press.
Reutskaja, E., & Hogarth, M. R. (2009). Satisfaction in Choice as a Function of the Number of Alternatives: When "Goods Satiate." *Psychology & Marketing*, *26*(3), 197–203. https://doi.org/10.1002/mar.20268
Schäfer, T., Sedlmeier, P., Städtler, C., & Huron, D. (2013). The psychological functions of music listening. *Frontiers in Psychology*, *4*, 1–33. https://doi.org/10.3389/fpsyg.2013.00511
Scheibehenne, B., Greifeneder, R., & Todd, P. M. (2009). What moderates the too-much-choice effect? *Psychology and Marketing*. https://doi.org/10.1002/mar.20271
Scheibehenne, B., Greifeneder, R., & Todd, P. M. (2010). Can there ever be too many options? A meta-analytic review of choice overload. *Journal of Consumer Research*, *37*(3), 409–425. https://doi.org/10.1086/651235
Schwartz, B. (2000). Self-determination: The tyranny of freedom. *American Psychologist*, *55*(1), 79–88. https://doi.org/10.1037/0003-066X.55.1.79
Schwartz, B. (2004). The Paradox of Choice: Why More is Less. In *Finance*. Harper Collins. http://wp.vcu.edu/univ200choice/wp-content/uploads/sites/5337/2015/01/The-Paradox-of-Choice-Barry-Schwartz.pdf%0Ahttp://books.google.com/books?hl=en&lr=&id=g422yyua-P8C&pgis=1
Schwartz, B., Ward, A., Lyubomirsky, S., Monterosso, J., White, K., & Lehman, D. R. (2002). Maximizing versus satisficing: Happiness is a matter of choice. *Journal of Personality and Social Psychology*, *83*(5), 1178–1197. https://doi.org/10.1037/0022-3514.83.5.1178
Sela, A., Berger, J., & Liu, W. (2009). Variety, vice, and virtue: How assortment size influences option choice. *Journal of Consumer Research*, *35*(6), 641–651. https://doi.org/10.1086/593692
Shah, A. M., & Wolford, G. (2007). Buying behavior as a function of parametric variation of number of choices. *Psychological Science*, *18*(5), 369–370. https://doi.org/10.1111/j.1467-9280.2007.01906.x
Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*,

69(1), 99–118. https://doi.org/10.2307/1884852
Simonson, I. (1990). The Effect of Purchase Quantity and Timing on Variety-Seeking Behavior. *Journal of Marketing Research*, *27*(5), 150–162.
Sloboda, J. A. (2010). Music in everyday life: The role of emotions. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (Issue January, pp. 493–514). Oxford University Press.
Spotify. (2020). *Spotify for Artists FAQ*. https://artists.spotify.com/faq/mastering-and-loudness#what-is-loudness-normalization-and-why-is-it-used
Spotify. (2021). *Spotify Company Info*. https://newsroom.spotify.com/company-info/
Statista. (2020). *Musik-Streaming*. https://de.statista.com/outlook/dmo/digitale-medien/digitale-musik/musik-streaming/weltweit
Steffens, J., Lepa, S., Egermann, H., Schönrock, A., & Herzog, M. (2017). Entwicklung eines Systems zur automatischen Musikempfehlung im Kontext des Audio Brandings. *Paper Presented at DAGA 2017 - Annual Conference of the German Acoustic Society, Kiel (Germany)*.
Taylor, S. E., & Brown, J. D. (1988). Illusion and Well-Being: A Social Psychological Perspective on Mental Health. *Psychological Bulletin*, *103*(2), 193–210. https://doi.org/10.1037/0033-2909.103.2.193
Timmermans, D. (1993). The impact of task complexity on information use in multi-attribute decision making. *Journal of Behavioral Decision Making*, *6*(2), 95–111. https://doi.org/10.1002/bdm.3960060203
Toffler, A. (1970). *Futureshock*. Random House.
Tversky, A., & Shafir, E. (1992). CHOICE UNDER CONFLICT:The Dynamics of Deferred Decision. *Psychological Science*, *3*(6), 358–361.
Van Herpen, E., & Pieters, R. (2002). The variety of an assortment: An extension to the attribute-based approach. *Marketing Science*, *21*(3), 331–341. https://doi.org/10.1287/mksc.21.3.331.144
Van Trijp, H. C. M. Van, Hoyer, W. D., & Inman, J. J. (1996). Why Switch? Product Category: Level Explanations for True Variety-Seeking Behavior. *Journal of Marketing Research*, *33*(8), 281–292. https://doi.org/10.2307/3152125
Walsh, J. W. (1995). Flexibility in Consumer Purchasing for Uncertain Future Tastes. *Marketing Science*, *14*(2), 148–165. https://doi.org/10.1287/mksc.14.2.148

Appendix

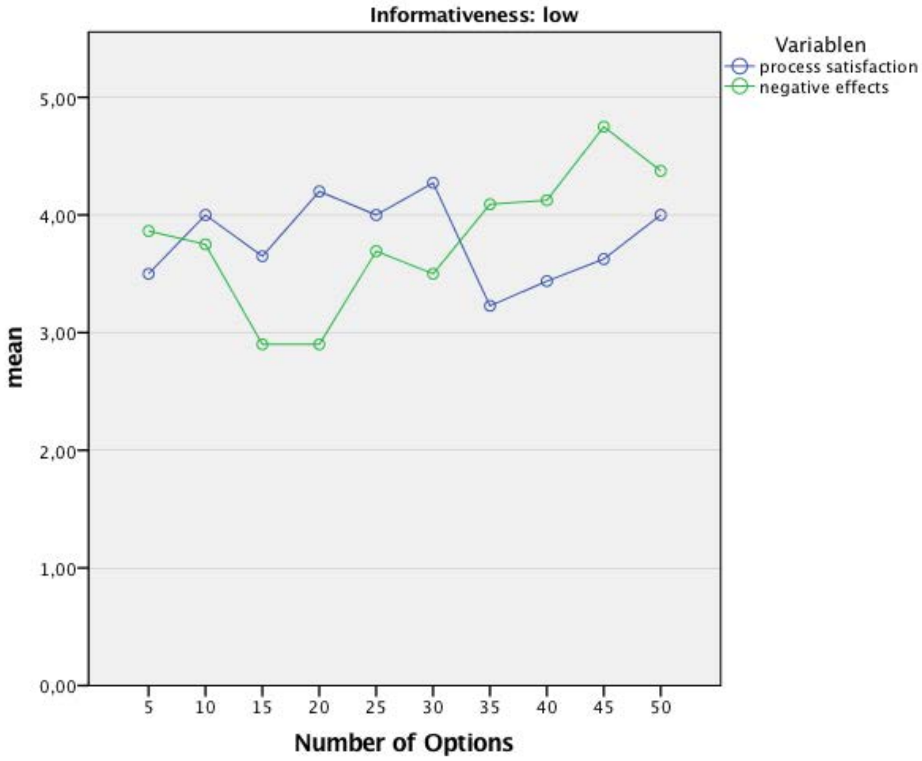


Figure 10. *Mean values for process satisfaction and negative effects across assortment size in the low informativeness condition.*

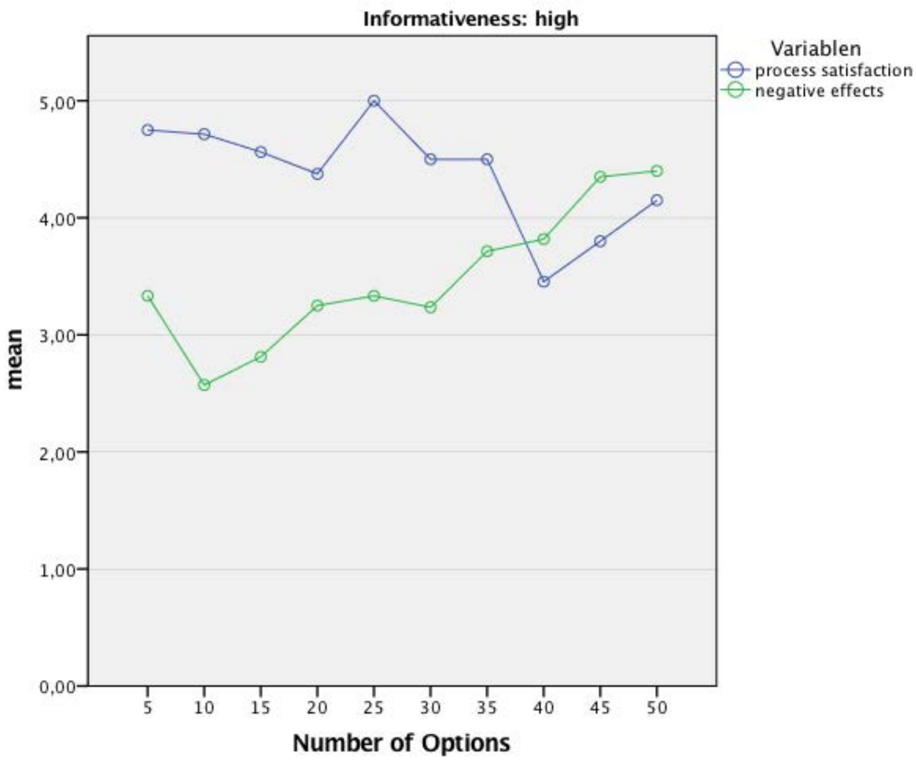


Figure 11. *Mean values for process satisfaction and negative effects across assortment size in the high informativeness condition.*

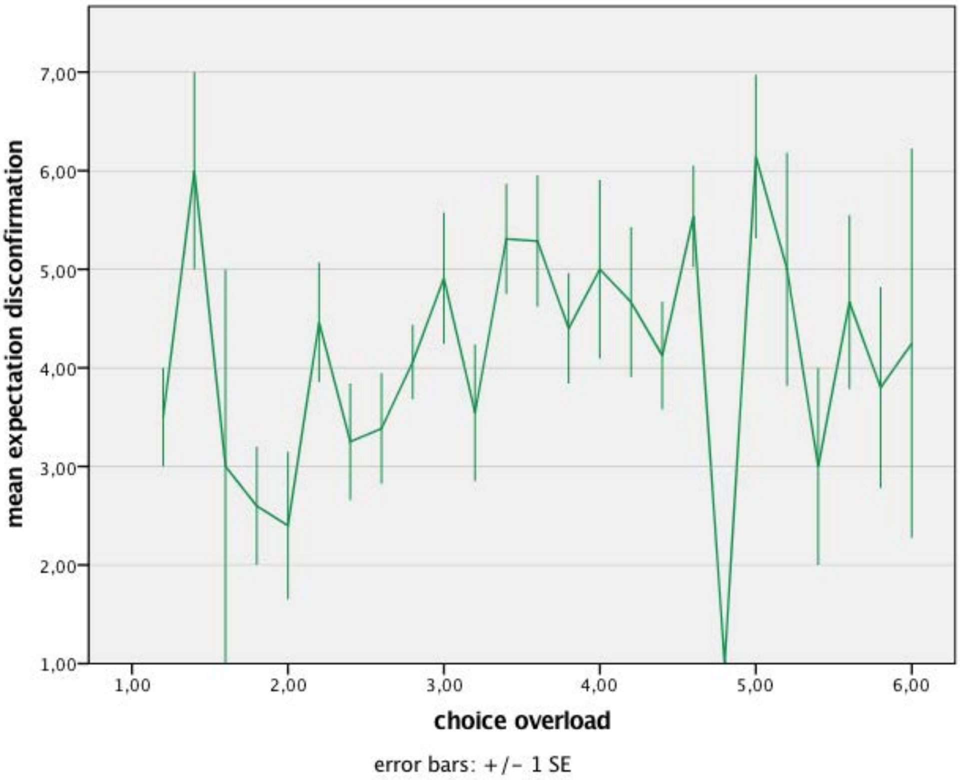


Figure 12. *Mean values for expectation disconfirmation across choice overload*

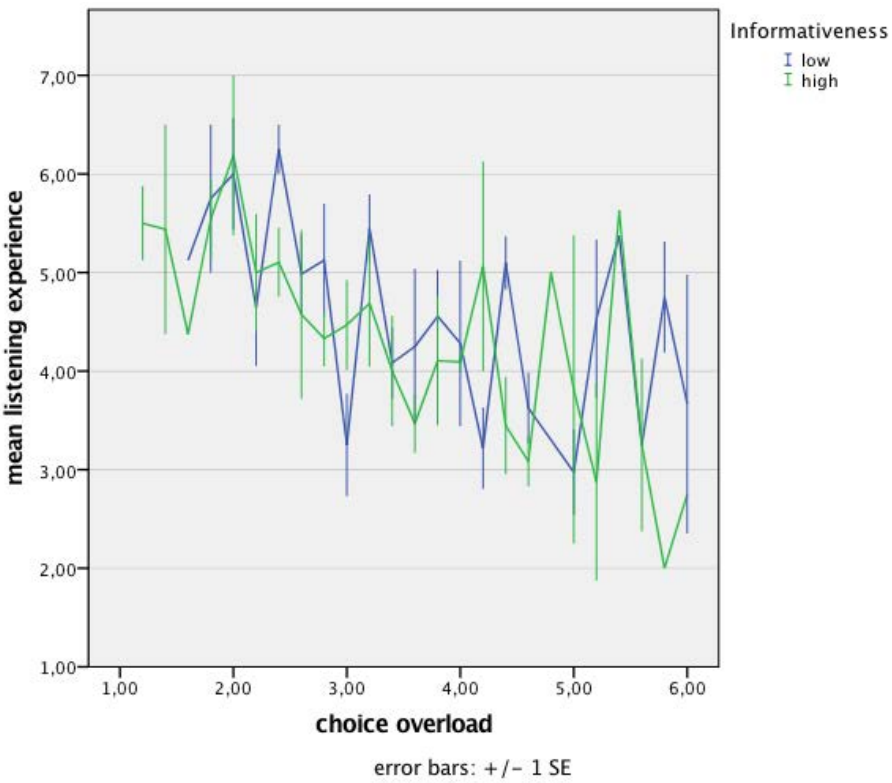


Figure 13. *Mean values for listening experience across choice overload for each informativeness condition.*