Exploring Online Music Listening Behaviors of Musically Sophisticated Users

Bruce Ferwerda

Department of Computer Science and Informatics
School of Engineering
Jönköping University
Jönköping, Sweden
bruce.ferwerda@ju.se

Marko Tkalčič Faculty of Computer Science Free University of Bozen-Bolzano Bozen-Bolzano, Italy marko.tkalcic@unibz.it

ABSTRACT

Due to the rise of available online music, a lot of music consumption is moving from traditional offline media to online sources. Online music sources offer almost an unlimited music collection to its users. Hence, how music is consumed by users (e.g., experts) may differ from traditional offline sources. In this work we explored how musically sophisticated users (i.e. experts) consume online music in terms of diversity. To analyze this, we gathered data from two different sources: Last.fm and Spotify. As expertise is defined by the ubiquitousness of experiences, we calculated different diversity measurements to explore how ubiquitous (in terms of diversity) the listening behaviors of users are. We found that different musical sophistication levels correspond to applying diversity related to specific kind of musical characteristics (i.e., artist or genre). Our results can provide knowledge on how systems should be designed to provide better support to expert users.

CCS CONCEPTS

• Information systems \rightarrow Recommender systems; • Human-centered computing \rightarrow User models; User studies.

KEYWORDS

Music; Expertise; Musical Sophistication; User Modeling; Music Listen Behaviors

ACM Reference Format:

Bruce Ferwerda and Marko Tkalčič. 2019. Exploring Online Music Listening Behaviors of Musically Sophisticated Users. In 27th Conference on User Modeling, Adaptation and Personalization Adjunct (UMAP'19 Adjunct), June 9–12, 2019, Larnaca, Cyprus, Jennifer B. Sartor, Theo D'Hondt, and Wolfgang De Meuter (Eds.). ACM, New York, NY, USA, Article 4, 5 pages. https://doi.org/10.1145/3314183.3324974

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

UMAP'19 Adjunct, June 9–12, 2019, Larnaca, Cyprus

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM. ACM ISBN 978-1-4503-6711-0/19/06...\$15.00 https://doi.org/10.1145/3314183.3324974

1 INTRODUCTION

Technologies are becoming increasingly pervasive and ubiquitous. Estimates indicate that currently almost 70% of the population owns some kind of smart device (e.g., smart phone, smart watch, tablet), with up to 90% in the more advanced markets [24]. With the increased penetration of technologies, these technologies can have a significant influence on how people shape their lives. Hence, technological influences could challenge our traditional knowledge of human behaviors, preferences, and needs as the possibilities that users have reach much further through the use of technologies. An example of how technologies can cause deviations from traditional knowledge is demonstrated by Ferwerda, Tkalčič, and Schedl [15]. They investigated the music listening behaviors of users of an online music streaming service and found that the listening patterns are richer than what was found among users listening to offline music collections (e.g., try out different content more and showed a more frequent change of taste). Others have investigated new relationships that have become available through new interaction opportunities that technologies facilitate. For example, personality traits are shown to relate to the way we browse (i.e., by genre, mood, or activity) for online music [17].

As user traits/characteristics play a fundamental role in segmenting users for user modeling purposes and its applications in personalization and adaptation of systems, it is important to understand how users manifest themselves in digital environments in support of their traits/characteristics. For example, personality has been extensively explored as a user trait for personalization purposes (see for an overview [19]). When it comes to user characteristics, domain expertise has often shown to play an influencing role; especially in fields as recommender systems. For example, Hu and Pu [20] showed that users with a higher degree of domain expertise in general find the recommendations that recommender systems provide to be less useful. A reason for this may be found in that recommender systems provide a list of recommendations to its users. Hence, there is a limited amount of options presented to users while online environments facilitate an abundance amount of content to choose from. Scheibehenne, Todd, and Greifeneder [28] showed in their meta-analytic review on choice overload that while more choices in general can cause negative effects, more choices have positive effects especially to those who are domain experts.

In this work we explored how self-reported musically sophisticated users (i.e., experts) manifest themselves through the abundance of content that online music streaming services have to offer, and thereby provide insights on how these users exploit the available online content to maintain their expertise. Moreover, we

analyzed different music characteristics (i.e., genre and artist) to explore on which level musically sophisticated users are focusing on. To explore the behavior of users of online streaming services, we collected data from two different sources: Last.fm ¹ and Spotify ² and let users fill in the Goldsmiths Musical Sophistication Index (Gold-MSI) questionnaire [25] to assess users' musical skills and behaviors on multiple dimensions. Our findings show that depending on the dimension of musical sophistication, users adopt different listening strategies. Our results provide indications of how musically sophisticated users use online music content to maintain their expertise as well as provide opportunities for online music streaming services to adapt their systems accordingly.

2 RELATED WORK

Identifying user traits/characteristics and understanding related behaviors is an important aspect for user modeling, which has implications on personalization and adaptation strategies of systems. User traits are often based on psychological models. A model that has received considerable attention is the use of personality traits. Specifically related to the music domain, an increasing amount of work has been done on identifying related behaviors to these traits (e.g., [8, 14, 31]) as well as work that has proposed to incorporate these traits into personalization and adaptation solutions (e.g., [3, 5, 7, 23]). Furthermore, attention has been given to implicitly acquire personality traits for personalization purposes by the use of (external) information sources, such as Facebook (e.g., [1, 10]), Twitter (e.g., [26]), Instagram (e.g., [9, 11–13, 22]), or a combination of such [30].

Alternatively to user traits, users can be segmented based on specific characteristics (e.g., age, gender, occupation). A characteristic that has often shown to be of importance is the level of expertise. The level of expertise has shown to influence how an individual experiences a system on factors such as usefulness, accuracy, and helpfulness (e.g., [20, 21]). A meta-analytic review on choice overload showed that experts benefit from more choices [28] instead of the restrictions that systems, such as recommender systems, provide them through compiled lists of items.

Expertise is trained through the ubiquitousness of experiences [2]. With technologies increasingly playing a central role in our lives, behaviors are increasingly moving online. Hence, the experiences that contribute to building up a level of expertise may happen through the use of technologies instead of through traditional offline experiences. In this work we explore how musically sophisticated users use online music streaming services in line with their expertise by analyzing music listening behaviors. The ubiquitousness of experiences, as how expertise is defined by, is assessed through the amount of diversity that is applied in music listening behaviors of users (i.e., more diverse music listening behavior implies more ubiquitous music listening). By analyzing online music listening behaviors, more specific diversity measurements can be assessed based on the characteristics of the music that users are listening to. In this case we took into account the artist and genre diversity of listening behaviors of users.

3 METHOD

Musical sophistication of users was assessed by using the Gold-MSI questionnaire developed by Müllensiefen et al. [25]. The Gold-MSI measures musical sophistication based on the following dimensions:

- Active engagement (how much time and money one spends on music)
- Perceptual abilities (cognitive musical ability related to music listening skills)
- Musical training (musical training and practice)
- Signing abilities (skills and activities related to singing)
- Emotions (active behaviors related to emotional responses to music)

Additionally, the Gold-MSI consists of a measurement of general musical sophistication, which incorporates aspects of the above mentioned sub-scales.

In order to investigate the behaviors of musically sophisticated users, we relied from data of two different sources: Last.fm (see Section 3.1) and Spotify (see Section 3.2). The participants in the Last.fm dataset were asked to fill in the general musical sophistication scale. A more in depth study was done by using Spotify: the participants in the Spotify dataset were asked to fill in specific sub-scales of the Gold-MSI. To keep the analyses manageable for both datasets, we only focus on the artist and genre characteristics of the music listening events.

To distinguish listening behaviors of musically sophisticated users from average users, we analyzed the amount of diversity in music listening behaviors. As users may listen to certain music more often than other music, measurements as uniqueness does not provide the full picture of diversity. Hence, we relied on diversity indices that are often used in ecology in which the number of different types are taken into account while also considering the distribution among these types. Common diversity indices to indicate biodiversity are the Shannon and the Simpson index, which are entropy-based indices. These diversity indices take into account the richness (the number of different types in a sample) and evenness (the relative abundance of different types making up the richness).

The Shannon index (H) is a diversity measurement that takes both richness and evenness of types in a given sample into account. Hence, the diversity entropy increases when richness as well as evenness increase. The Shannon index is calculated by the following formula where R is the richness of the sample and p_i is the evenness of the i^{th} type:

$$H = -\sum_{i=1}^{R} p_i \log p_i$$

The Simpson index (D) is a diversity measurement based on the dominance of types within a sample. It puts more weight to common or dominant types, and therefore the calculated diversity entropy is less susceptible to be influenced by types consisting with only few representatives. The Simpson index is calculated by the following formula where R is the richness of the sample and p_i is the evenness of the i^{th} type:

$$D = \frac{1}{\sum_{i=1}^{R} p_i^2}$$

¹https://www.last.fm/

²https://www.spotify.com/

3.1 Last.fm

We recruited 254 participants through Amazon Mechanical Turk for the first part of the study. Participation was restricted to those located in the United States with a very good reputation (\geq 95% HIT approval rate and \geq 1000 HITs approved) and a Last.fm account with at least 25 listening events. Furthermore, participants were asked to fill in the questions of the general musical sophistication factor. A compensation of \$1 was provided.

We crawled the complete listening history of each participant, which resulted in a dataset consisting of 13,624,84 tracks. We then aggregated the listening events to represent artist and playcount (i.e., number of times listened to an artist). Through the LastFM API, we crawled additional information about the artists by using the "Artist.getTopTags" endpoint. This endpoint provided us with all the tags that users assigned to an artist, such as instruments ("guitar"), epochs ("80s"), places ("Chicago"), languages ("Swedish"), and personal opinions ("seen live" or "my favorite"). Tags that encode genre or style information were filtered for each artist. The filtered tags were indexed by a dictionary of 18 genre names retrieved from Allmusic. ³ For each user, the artists that were listened to were aggregated by the indexed genre with their play-count.

3.2 Spotify

A total of 61 participants were recruited through a participant pool managed by the Human-Technology Interaction group at Eindhoven University of Technology. Using Spotify's API we retrieved the participants' top tracks, which resulted in a dataset consisting of 21.080 tracks.

The Spotify data was collected as part of a larger study that investigated how the order of individual songs affects playlist experience. To study the relationship between musical sophistication and music listening behavior, participants' logged into our app through the Spotify API, which allowed us to retrieve their music listening behavior. In addition, participants completed a survey with items from the Goldsmiths Musical Sophistication Index (Gold-MSI; [25]). The Gold-MSI measures musical sophistication on five subscales: active engagement, emotions, singing abilities, perceptual abilities, and musical training. However, in this work, we only asked participants to respond to two of these subscales as we believe that they are the most prominent ones reflected in online music listening behaviors:

- (1) Active engagement (measured by 9-items).
- (2) Emotions (measured by 6-items).

4 RESULTS

Spearman's correlation was computed between the different musical sophistication factors and the calculated diversity entropy to assess the relationship of levels of sophistication and music listening behaviors. The reported significant results adhere to alpha levels of p < .05 (see Table 1). Whereas the Shannon diversity index is connected to the data richness, the Simpson diversity index is connected to the evenness of the data.

Results indicate that based on the general musical sophistication measurement, participants who reported to have a higher degree of

| Last.fm | General Musical Sophistication | | | |
|---------|--------------------------------|---------|----------|----------|
| | Shannon | Simpson | Richness | Evenness |
| Artist | 0.166 | 0.103 | 0.183 | -0.048 |
| Genre | 0.191 | 0.143 | 0.159 | 0.124 |
| | | | | |
| Spotify | Active Engagement | | | |
| | Shannon | Simpson | Richness | Evenness |
| Artist | -0.234 | -0.217 | -0.178 | -0.199 |
| Genre | -0.035 | 0.023 | -0.091 | 0.102 |
| | | | | |
| Spotify | Emotions | | | |
| | Shannon | Simpson | Richness | Evenness |
| Artist | -0.164 | -0.203 | -0.095 | -0.236 |
| Genre | 0.158 | 0.225 | 0.025 | 0.253 |

Table 1: Spearman's correlation between musical sophistication factors and diversity (related) variables: Shannon and Simpson diversity indices, richness, and evenness. Bold faced numbers adhere to a p < 0.05 significance level.

general musical sophistication, show more diverse listening behaviors on an artist (rs(254) = 0.166, p = 0.043) as well as on a genre (rs(254) = 0.191, p = 0.038) level.

When it comes to the specific musical sophistication levels, participants who indicated to have a higher degree of active engagement, show less diverse behavior on an artist level than others. This is reflected in negative correlations for both Shannon (rs(61) = -0.234, p = 0.049) and Simpson (rs(61) = -0.217, p = 0.042) diversity indices.

The participants scoring higher on the emotions scale of the musical sophistication index show a more diverse pattern in their listening behaviors. On an artist level a negative correlation was found (rs(61) = -0.203, p = 0.039), while on a genre level a positive correlation was found (rs(61) = 0.225, p = 0.041). The results indicate that those scoring higher on the emotions scale tend to apply less diversity on an artist level in their listening behaviors, but more diversity on a genre level.

5 DISCUSSION

Our results show that there are significant differences in online music listening behaviors between the average listener and those who self-reportedly indicate to be on a more sophisticated level (i.e., experts). On a general level of musical sophistication, our results show that higher amount of musical sophistication is correlated with a higher degree of applying diversity to listening behaviors. This higher degree of diversity is reflected on both an artist and a genre level.

When looking at more specific factors of musical sophistication, we found more diverging results. The active engagement factor showed to be negatively correlated with artist diversity and no significant correlation with genre diversity at all. Likewise, the emotions factor showed a negative correlation on an artist level, but a positive correlation on a genre level.

The positive correlations are in line with expectations on how experts would keep up with their level of expertise: to explore and train their knowledge. In this case the exploration would come

³http://www.allmusic.com

from diversification in music listening behaviors. The negative correlation of artist diversity against the positive correlation of genre diversity within the emotions factor might be explained by the characteristics of this kind of sophistication. The emotions factor depicts sophistication behaviors such as the ability to select music that evoke the right emotions. Hence, suitable emotions of music may be more facilitated through different genres than through different artists.

More remarkable are the correlation results of the active engagement factor. The active engagement factor depicts sophistication activities that are related to engagement behaviors (e.g., keeping track of the music one comes across) and allocating time and money (e.g., listening to music for an X amount of hours per day). Hence, it may well be that the diversification of these music listeners is expressed on a song level rather than on an artist or genre level. The negative correlation on an artist level may then be explained by the fact that these users may listen to more diverse songs from a select number of artists (i.e., less diverse artists). This is confirmed through the negative correlations of richness as well as evenness, meaning that the listening events of these kind of sophisticated users in general are less rich (i.e., fewer artists listened to) and less evenly (i.e., certain artists get more attention) distributed.

6 CONCLUSION, IMPLICATIONS

In this work we investigated how self-reported musically sophisticated music listeners (i.e., experts) manifest themselves through the abundance of online music that is available today. We collected data from two different online streaming services (i.e., Last.fm and Spotify) to investigate how these users exploit online music content to maintain their expertise. By investigating the applied diversity on different levels of music characteristics (i.e., artist and genre level), we found that on a general level musically sophisticated users apply more diversity to their listening behaviors on both an artist and genre level. However, when considering more specific musical sophistication levels, the level of applied music diversity differ. Those who are more musically sophisticated through active engagement with music did not show an increased applied music diversity, and even showed less diversified behavior on an artist level. It may well be that diversity related to active engagement expresses itself through the diversity of songs instead of through artist and genre. Furthermore, musically sophisticated participants on an emotions level are related to behaviors that engage in finding and recognizing music that fit certain emotions. Hence, diversified music listening behaviors are more expressed through more diverse genres and less to that of artists. These findings are in line with previous research that showed that the effects of expertise is not universal, but depending on the domain of the expertise [18].

Our work contributes to an understanding on how technologies are influencing and shaping our lives. By gaining a better understanding of how experts are behaving to maintain their level of expertise, systems can be developed to better support their needs. Prior work has shown that implicit acquisition of musical sophistication levels is possible [4]. Hence, being able to detect the musical sophistication of users could provide personalized systems the opportunity to adapt the interface to better support the diversification needs of these users.

The work presented in this paper has a couple of limitations. To keep the analyses manageable, we decided to only focus on diversity on a artist and genre level. By not analyzing song level diversity, we have limited ourselves to fully explain the results of some of the found effects (or no effects). For example, the results of the active engagement factor may have been able to be further explained through song diversity. Furthermore, we only used two subscales of Gold-MSI. Future work should address the inclusion of the other three subscales as well. This will be done in conjunction with plans to expand our dataset by increasing the number of participants in our Spotify dataset and the number of measurements per participant. It will allow for a more in-depth investigation of the relationship between behavioral features and the Gold-MSI scores. Lastly, prior work has shown that cultural differences exist in music consumption [6, 16, 27, 29]. This should be taken into account in future work as well.

7 ACKNOWLEDGEMENTS

We would like to thank Dr. Mark Graus from Maastricht University in the Netherlands for sharing his Spotify dataset with us.

REFERENCES

- Fabio Celli, Elia Bruni, and Bruno Lepri. 2014. Automatic personality and interaction style recognition from facebook profile pictures. In Proceedings of the 22nd ACM international conference on Multimedia. ACM. 1101–1104.
- [2] Harry Collins and Robert Evans. 2008. Rethinking expertise. University of Chicago Press.
- [3] Bruce Ferwerda. 2016. Improving the User Experience of Music Recommender Systems Through Personality and Cultural Information. PhD. Johannes Kepler University Linz, Linz, Austria (2016).
- [4] Bruce Ferwerda and Mark Graus. 2018. Predicting Musical Sophistication from Music Listening Behaviors: A Preliminary Study. arXiv preprint arXiv:1808.07314 (2018).
- [5] Bruce Ferwerda and Markus Schedl. 2014. Enhancing Music Recommender Systems with Personality Information and Emotional States: A Proposal.. In Umap workshops.
- [6] Bruce Ferwerda and Markus Schedl. 2016. Investigating the Relationship Between Diversity in Music Consumption Behavior and Cultural Dimensions: A Cross-Country Analysis.. In UMAP (Extended Proceedings).
- [7] Bruce Ferwerda and Markus Schedl. 2016. Personality-based user modeling for music recommender systems. In Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer, 254–257.
- [8] Bruce Ferwerda, Markus Schedl, and Marko Tkalcic. 2015. Personality & emotional states: Understanding users' music listening needs. UMAP.
- [9] Bruce Ferwerda, Markus Schedl, and Marko Tkalcic. 2015. Predicting personality traits with instagram pictures. In Proceedings of the 3rd Workshop on Emotions and Personality in Personalized Systems 2015. ACM, 7–10.
- [10] Bruce Ferwerda, Markus Schedl, and Marko Tkalcic. 2016. Personality traits and the relationship with (non-) disclosure behavior on facebook. In *Proceedings of the 25th International Conference Companion on World Wide Web*. International World Wide Web Conferences Steering Committee, 565–568.
- [11] Bruce Ferwerda, Markus Schedl, and Marko Tkalcic. 2016. Using instagram picture features to predict users' personality. In *International Conference on Multimedia Modeling*. Springer, 850–861.
- [12] Bruce Ferwerda and Marko Tkalcic. 2018. Predicting Users' Personality from Instagram Pictures: Using Visual and/or Content Features?. In Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization. ACM, 157–161.
- [13] Bruce Ferwerda and Marko Tkalcic. 2018. You Are What You Post: What the Content of Instagram Pictures Tells About Users' Personality. In The 23rd International on Intelligent User Interfaces, March 7-11, Tokyo, Japan.
- [14] Bruce Ferwerda, Marko Tkalcic, and Markus Schedl. 2017. Personality traits and music genre preferences: how music taste varies over age groups. In 1st Workshop on Temporal Reasoning in Recommender Systems (RecTemp) at the 11th ACM Conference on Recommender Systems, Como, August 31, 2017., Vol. 1922. ACM Digital Library, 16–20.
- [15] Bruce Ferwerda, Marko Tkalcic, and Markus Schedl. 2017. Personality Traits and Music Genres: What Do People Prefer to Listen To?. In Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization. ACM, 285–288.
- [16] Bruce Ferwerda, Andreu Vall, Marko Tkalcic, and Markus Schedl. 2016. Exploring music diversity needs across countries. In Proceedings of the 2016 Conference on

- User Modeling Adaptation and Personalization. ACM, 287-288.
- [17] Bruce Ferwerda, Emily Yang, Markus Schedl, and Marko Tkalčič. 2015. Personality traits predict music taxonomy preferences. In Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems. ACM, 2241–2246.
- [18] Bruce Ferwerda, Emily Yang, Markus Schedl, and Marko Tkalcic. 2019. Personality and taxonomy preferences, and the influence of category choice on the user experience for music streaming services. *Multimedia Tools and Applications* (27 Feb 2019). https://doi.org/10.1007/s11042-019-7336-7
- [19] Mark Graus and Bruce Ferwerda. 2019. Theory-Grounded User Modeling for Personalized HCI. DeGruyter Oldenbourg. In Press.
- [20] Rong Hu and Pearl Pu. 2010. A study on user perception of personality-based recommender systems. In *User modeling, adaptation, and personalization*. Springer, 291–302
- [21] Arnold Kamis and Michael J Davern. 2004. Personalizing to product category knowledge: exploring the mediating effect of shopping tools on decision confidence. In null. IEEE, 70202a.
- [22] Alixe Lay and Bruce Ferwerda. 2018. Predicting Users' Personality Based on Their 'Liked' Images on Instagram. In The 23rd International on Intelligent User Interfaces, March 7-11, 2018.
- [23] Michael J Lee and Bruce Ferwerda. 2017. Personalizing online educational tools. In Proceedings of the 2017 ACM Workshop on Theory-Informed User Modeling for Tailoring and Personalizing Interfaces. ACM, 27–30.
- [24] Zenith Media. 2018. Mobile Penetration. (2018). https://www.zenithmedia.com/ smartphone-penetration-reach-66-2018/ date accessed: 19 November 2018.

- [25] Daniel Müllensiefen, Bruno Gingras, Jason Musil, and Lauren Stewart. 2014. The musicality of non-musicians: an index for assessing musical sophistication in the general population. PloS one 9, 2 (2014), e89642.
- [26] Daniele Quercia, Michal Kosinski, David Stillwell, and Jon Crowcroft. 2011. Our twitter profiles, our selves: Predicting personality with twitter. In 2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing. IEEE, 180–185.
- [27] Markus Schedl, Florian Lemmerich, Bruce Ferwerda, Marcin Skowron, and Peter Knees. 2017. Indicators of country similarity in terms of music taste, cultural, and socio-economic factors. In 2017 IEEE International Symposium on Multimedia (ISM). IEEE, 308–311.
- [28] Benjamin Scheibehenne, Rainer Greifeneder, and Peter M Todd. 2010. Can there ever be too many options? A meta-analytic review of choice overload. *Journal of Consumer Research* 37, 3 (2010), 409–425.
- [29] Marcin Skowron, Florian Lemmerich, Bruce Ferwerda, and Markus Schedl. 2017. Predicting genre preferences from cultural and socio-economic factors for music retrieval. In European Conference on Information Retrieval. Springer, 561–567.
- [30] Marcin Skowron, Marko Tkalčič, Bruce Ferwerda, and Markus Schedl. 2016. Fusing social media cues: personality prediction from twitter and instagram. In Proceedings of the 25th international conference companion on world wide web. International World Wide Web Conferences Steering Committee, 107–108.
- [31] Marko Tkalčič, Bruce Ferwerda, David Hauger, and Markus Schedl. 2015. Personality correlates for digital concert program notes. In *International Conference on User Modeling, Adaptation, and Personalization*. Springer, 364–369.