

What is fair? Exploring the artists' perspective on the fairness of music streaming platforms

Andrés Ferraro¹, Xavier Serra¹, and Christine Bauer²

¹Universitat Pompeu Fabra, andres.ferraro@upf.edu

²Utrecht University, c.bauer@uu.nl

This is the accepted version of the paper appearing in the Proceedings of the 18th IFIP International Conference on Human-Computer Interaction (INTERACT 2021).

Andrés Ferraro, Xavier Serra and Christine Bauer. "What is fair? Exploring the artists' perspective on the fairness of music streaming platform". 18th IFIP International Conference on Human-Computer Interaction (INTERACT 2021). Bari, Italy.

Abstract

Music streaming platforms are currently among the main sources of music consumption, and the embedded recommender systems significantly influence what the users consume. There is an increasing interest to ensure that those platforms and systems are fair. Yet, we first need to understand what fairness means in such a context. Although artists are the main content providers for music platforms, there is a research gap concerning the artists' perspective. To fill this gap, we conducted interviews with music artists to understand how they are affected by current platforms and what improvements they deem necessary. Using a Qualitative Content Analysis, we identify the aspects that the artists consider relevant for fair platforms. In this paper, we discuss the following aspects derived from the interviews: fragmented presentation, reaching an audience, transparency, influencing users' listening behavior, popularity bias, artists' repertoire size, quotas for local music, gender balance, and new music. For some topics, our findings do not indicate a clear direction about the best way how music platforms should act and function; for other topics, though, there is a clear consensus among our interviewees: for these, the artists have a clear idea of the actions that should be taken so that music platforms will be fair also for the artists.

1 Introduction

Music streaming platforms are currently among the main sources of music consumption. What the users consume is strongly influenced by what is offered on the music platforms, and what is promoted through the algorithmic recommendations in particular. What the users consume, in turn, shapes the music streaming ecosystem at large. There is an increasing research interest in making the platforms and their recommendations fairer.

Generally, the topic of fairness received particular attention in machine learning and artificial intelligence [35], focusing on the quantification of fairness and the algorithmic perspective. Research in the field of human-computer interaction (HCI), in contrast, studied the perception of fairness by taking the end consumers' perspective (e.g., [31, 30, 54]). In general, a large part of HCI research addresses the user who directly interacts with a (computational) system. Yet, also other humans are impacted by system design and HCI practices. Hence, it is crucial to consider also these groups' demands in design considerations.

Recommender systems connect and affect several stakeholders including consumers, the platform provider, item providers, and society [36]. While several potential sources of bias have been identified that may lead to unfairness (e.g., [57, 40, 22]), the focus of attention lies on algorithmic

and data bias [8]. A recent review [43] highlights that existing literature tends to consider the implications for the receivers of the recommendations (i.e., end consumers) only. Thus, the research gap lies in considering the provider's interests (and the interests of the society at large) when assessing the ethical impact of recommender systems.

Addressing this research gap, this work focuses on the item provider's perspective on online platforms, which are an underrepresented group of humans in HCI research. This also holds for the music domain. Recommender systems play a prominent role in connecting users with content (and the artists behind this content) on such music platforms. While various studies (e.g., [12, 37]) have investigated and shown that current music recommender systems (MRS) do not serve different user segments equally well, there is a research gap concerning the artists' perspective [9, 10]. Besides being a so-far underrepresented group of humans in HCI research, music artists have limited opportunities for direct interaction with a music platform (in their role as an item provider)¹; and yet, they are strongly affected by the design of such systems. While our long-term goal is to have music platforms that are perceived 'fair' from the various stakeholders' perspectives, this work zooms in music artists as the main item providers for music platforms. The concrete goal of this work at hand is to *understand the different dimensions of music platforms (and their recommender systems) that define fairness from the artists' perspective*.

To this end, we conducted semi-structured interviews with music artists from different countries and with varying popularity. Using a Qualitative Content Analysis as proposed by [41], we have (i) identified various aspects that show how the artists feel affected by the current music platforms and their integrated information retrieval and recommender systems components, and (ii) how such music platforms and systems could be improved concerning fairness. Only such understanding can ultimately lead to fairer music platforms that are not only optimized for the interests of the platform providers or the consumers. We deem our work a fundamental contribution to (i) understanding the impact that such music platforms have on artists from their perspective and (ii) understanding what the artists consider fair for them.

The remainder of this work is structured as follows: First, we present the conceptual basis and discuss related work (Section 2). In Section 3, we detail the methods. Section 4 represents the major part of this paper, where we outline and discuss the issues and aspects that we derived from the interviews. As we illustrate the artists' concerns and ideas, we embody a completely novel contribution to the field. In Section 5, we discuss how the findings reflect the existing understanding of fairness and derive implications for the operationalization of recommender systems and interaction functionalities on music platforms. In Section 6, we summarize our work and point to future research directions.

2 Conceptual Basis and Related Work

A purely technical approach is not sufficient for defining and operationalizing fairness in practice, as seen in previous attempts in several fields [19, 18, 39, 33]. Taking an information systems perspective, [27] emphasize that fair systems need to embrace people, technology, and organizations, and that unsolved challenges exist on all three dimensions. In other words, fairness cannot be operationalized with a lack of a definition or understanding of what fairness connotes in a certain context. The notion of fairness has evolved over time [35]; and particularly recent literature (e.g., [33, 48]) emphasizes that for developing a good understanding of fairness in a given context, it is crucial to take a multidisciplinary point of view and listen to the opinions of the various stakeholders involved and affected.

2.1 Defining Algorithmic Fairness

There is a multitude of definitions of (algorithmic) fairness [35]. Two common definitions are 'group' and 'individual' fairness. Individual fairness reflects that similar individuals should be treated similarly. Group fairness ensures that people of a protected group should be treated in the same way as the rest of the population. While [21] clearly distinguishes those two concepts, other works (e.g., [15]) suggest that individual and group fairness are not contradictory and may be achieved simultaneously. In information retrieval, fairness is frequently defined in terms of *exposure* [14, 47] or *attention* [49]. Thereby, [14] focus on individual fairness, whereas other works typically consider group fairness [49, 47]. The idea of exposure of the artists could also be adopted

¹Interaction is limited to providing recordings and some meta-data (e.g., title, tags).

for the music domain. Yet, it is not answered how a *fair exposure* should be operationalized or—on an even deeper level—what fairness actually means in the context of music platforms. To understand what fairness is from the artists' perspective, we need to involve artists. To the best of our knowledge, there is no public work that reaches out to artists to identify how they feel affected by current music platforms and how they believe the embedded recommenders systems could be improved to be fair for them.

2.2 Perception of Fairness in Algorithms

Besides defining algorithmic fairness, there is a wealth of studies investigating how people perceive and reason about the fairness of algorithms. When it comes to definition, people seem to prefer simple compared to complex ones [51].

In a survey investigating the general perception of fairness of automated decision-making systems, [31] found that people consider that systems make more objective decisions than humans and may also process larger amounts of information which would allow them to make fairer decisions. Yet, the respondents also noted that systems are limited concerning generalizability and modeling reality. Another survey [30] studied how people perceive the fairness of realistically-imperfect systems. As the respondents had contradictory opinions on systems, the authors conclude that it is impossible to achieve a broad acceptance in society regarding the “right” fairness definition. The study also highlights that there is a general preference towards human judges compared to a system, even if the participants consider the systems fair or unbiased.

In an online experiment [54], participants rated algorithms according to their perception of fairness. The study found that participants rate systems as more fair if they are favored by it, even in cases where the algorithms were explicitly described to the participants as being very biased to a particular demographic group. This effect was present across various participant groups, in different levels depending on education level, gender, and other aspects of the participants.

In group discussions and interviews with traditionally marginalized groups in the US, [56] sought to understand these groups' perceptions of fairness in algorithms. This work highlights that the opinions regarding the fairness of algorithms vary depending on individual factors, context, different stakeholders' perspectives, and different framing of fairness. This can help to explain inconsistencies in some findings across studies on fairness perception. Further, this shows that the contextual factors should be taken into account when studying algorithmic fairness, supporting the idea of considering the interests of the different groups of stakeholders of the systems.

2.3 Bias and Fairness in Music Recommendation

Besides the general issues and considerations concerning fairness, the music platforms and their recommenders face specific challenges. Among others, the inherent popularity bias is widely discussed. In this section, we discuss related works that study how music recommender can affect artists.

The music market is a typical example of the long-tail economy [6, 7] with a highly uneven distribution of demand for the most popular and the least popular items. Since Anderson [6, 7] popularized the concept, there is an ongoing debate about whether online platforms indeed facilitate users to consume items in the long tail or, on the contrary, accentuate the ‘superstar phenomenon’ [46] where the popularity curve gets more skewed in the long-term [17, 11, 24]. Celma [16] was among the first to study how different recommenders may or may not promote the less popular items in the music domain. Also for other domains, [28] found that users tend to consume a reduced number of items when recommender systems are at play. While these works focused on the consumers' behavior, the findings clearly indicate that recommenders have an impact on item providers.

A study in the Spotify context [5] found that users who follow algorithmically-generated recommendations have reduced diversity in the content they consume. In contrast, users who consume increasingly more diverse content, are those who reduce the algorithmic consumption and increase their organic consumption. As reducing the diversity in recommendations means that fewer artists get exposure and that it is more difficult for them to reach an audience, we believe that is important to also study the effect that algorithmically-generated recommendations can have on artists (e.g., [23, 25].)

Based on the idea that many local artists tend to be obscure long-tail artists, there are research endeavors (e.g., [52, 3]) to “localify” recommendations to promote local artists. Although technical

approaches to localize music recommendations are proposed, the artists' perspective on such a goal or approach is not discussed. Another work by Spotify [55] studies music consumption on a country level. The latter study investigates how users from a country consume content from another country, and how this consumption pattern evolves and changes over time. The authors identify that language and geographical proximity potentially impact the consumption between countries. Yet, a more fine-grained investigation is needed where also the popularity of the artist is taken into account. Such an endeavor may allow drawing conclusions about the chances given to local artists for gaining a larger audience using these platforms.

Mehrotra et al. [42] acknowledge that fair music recommenders need to consider multiple stakeholders. They define the recommendation problem as a marketplace and consider the perspectives of the music consumers and the artists. Intending to achieve algorithmic fairness, they define a fairness metric for artists that they base on the popularity distribution in the recommendations. While that work points out that there are different ways of defining fairness, a rationale for their choice and definition are missing.

However, previous work rarely reaches out to artists and, to date, no work involves music artists for understanding how music platforms can be fair for them. Baym [13] builds on interviews with artists to investigate how digital technologies changed their interaction with the audience. Andersen and Knees [4] interview artists for investigating their assessment of algorithms and artificial intelligence in the music creation process; our work, in contrast, refers to recommenders for music consumers. Aguiar and colleagues [2, 1] analyze how strategic decisions and recommenders embedded in music platforms affect music consumption behavior and the success of songs. Ferraro and colleagues [23, 25] take an algorithmic approach to analyze the impact of MRS on users' behavior and how this affects different groups of artists. Holzapfel et al. [34] raise ethical implications in music information retrieval and conceptually illustrate how music platforms may negatively impact certain groups of artists.

3 Methods

To understand the artists' perspective on the current music platforms and their embedded MRS, we conducted semi-structured interviews with music artists. The interviews took place from December 2019 through March 2020. The interviews are part of our ongoing research to explore viable solutions for fairer music platforms and MRS, this work at hand focuses on understanding the artists' perspective on what represents a fair music platform.

We carried out 9 semi-structured interviews with music artists. According to research practice [20, 44], the sample size is adequate and, more importantly, we reached a high level of thematic saturation [29] with the same topics being repeatedly mentioned across the interviews. The interviews were designed to last one hour in total for each interview including a brief and general introduction to MRS and music platforms. We used open questions with the aspiration that the artists would bring up their own ideas that might not be considered in the field before. In addition, we proposed specific alternatives to gather the artists' opinions on those alternatives.

In the following, we detail the interview process. The materials used for the interviews (i.e., invitation letter, consent form, first version of questions, final version of questions) can be found on Zenodo².

3.1 Interviews

Before the interviews, we informed that the data and results of the interviews will be kept anonymous at all times, which was important so the artists could feel free to share concerns regarding the music platforms or issues related to the music industry. In addition to the consent form, we asked the participants to fill out a short form with optional information about themselves that would be used to refer to their answers (i.e., age, gender, country, genre, popularity level, number of records/singles, years active in the music industry, contracts with labels). The first 10 minutes of the interview were used to explain the project's purpose and to give a general introduction to MRS and how these are integrated into current music platforms.

For the interview part, we defined a protocol to be used during the interviews which included a set of guiding questions and tentative question to encourage the interviewees to elaborate further. The protocol was developed so that each guiding question addressed and explored various

²<https://doi.org/10.5281/zenodo.4793395>

topics and issues of how the music platforms might affect artists. As the first step in elaborating this guide, we started with collecting information and ideas about the general aspects of current music platforms that could affect artists. In a second step, we formulated questions to address the identified issues. Since the rich collection of potentially interesting issues had to be reduced to fit the time scope of an interview and to have (a narrower) focus, in the third step, each team member gave a score between 1 and 3 to each of the original 36 questions according to priority, with 1 being the highest priority. Afterward, the team discussed the 14 questions with differing scores until consensus was reached. Based on the agreement between the scores we defined a definitive list of guiding 21 questions (whereof we use 11 as main questions and 10 as sub-questions) to be used in the interview protocol (Table 1). Note, all interviews were held in Spanish and the materials provided to participants were all in Spanish.

Table 1: Guiding questions in the interview protocol.

No.	Topic	Guiding question
1	Convey interest, gain trust	Do you use any platform to listen to music? What's your experience with it?
2	Reflecting	Do you think your career would be much different without these systems?
3	Lack of control	Which of your music tracks should be recommended more and which less?
4	Bias to more popular	There are groups of artists that are not recommended by the system because of different reasons. Do you see any alternatives for this?
5	Diversity	Music considered "niche music" is not recommended to many users, should the system nurture diversity (e.g., in terms of genres, styles, artists from all over the world, popular and not-yet-popular) or focus more on recommending what the user is familiar with?
6	Size of repertoire	If artist X has more music than the artists Y, do you think the system should recommend more music by artist X—or should the recommendation be independent of an artist's repertoire?
7	New artists	For a given user out of 100 recommendations, how many do you think should be new artists?
8	New music	Should your older songs be promoted more than your newer songs?
9	Country quotas	In a music platform that has more users from country X but more artists from country Y, the artists from X could be recommended more than artists from Y. What is the behavior that you expect from the system in that case?
10	Influencing the users	Currently, K-pop is the 7th most listened genre, over R&B and classical music. Such a popularity distribution could also refer to gender, country, or other aspects of the music. Do you think the systems should try to reproduce this behavior, or should try to provoke a change on it?
11	Income distribution	Do you think the current model based on number of streams is good or there could be a better model?

We note limitations of our interview design. First, while we cater for diversity, our sample is not representative of the population of music artists. The findings we report should, thus, be viewed as a deep exploration of our sample's beliefs and attitudes, but not as generalizing to the artist population as a whole. Yet, the findings indicate that we reached a high level of saturation with the same topics being repeatedly mentioned across the interviews. Second, while we assure anonymity in the presentation of the results, the interviewer naturally knows the participants' identity, which may have influenced participants such that some issues may not have been voiced. As some of the participants sometimes used strong language and addressed delicate issues, we believe that we were able to maintain a comfortable atmosphere with a high level of trust.

3.2 Participants

We recruited 9 music artists that we consider diverse—in the kind of music they perform, their popularity, location, age, and gender. Four of the participants are between 26–35 years old, four are between 36–45, and one is within the range 46–55. Seven are male and two are female. Most of these artists have many projects in parallel (one is a solo artist, the others work with many bands). The artists were born in four different countries (i.e., Australia, Spain, Russia, Uruguay) and started their careers in three different countries (i.e., Spain, Uruguay, Cuba).

The genres that the artists consider their music sum up to a total of 24. Examples are folk, pop, ska, punk-rock, dubstep, D&B, and world music. The number of years in the industry is between 4 and 25. The number of albums released ranges between one and ten. Five participants consider themselves ‘independent artist’, three have a contract with one of the three major labels (i.e., Sony, Universal, Warner), two have a contract with an independent label. Five artists consider themselves internationally known, four are known within their country. The specific information of each participant is given in Table 2, including the identifier that we use to refer to participants for quotes.

Table 2: Information about the participants.

ID	Country	Age	Music Styles	Audience	Gender	Contract
P1	Uruguay	46-55	Rock, Folk, Hip-Hop, Electronic	International	Male	Major Label
P2	Uruguay	26-35	Rock, Hip-Hop, Reggae, Dub	Local	Male	Major Label
P3	Uruguay	36-45	Ska, Punk/Rock, Dub, Dubstep, D&B	International	Male	Major Label
P4	Spain	26-35	Indie, Rock	Local	Male	Independent
PF1	Cuba	36-45	World, Jazz, Cuban Music, Electronic	International	Female	Independent
PF2	Spain	26-35	Indie Pop, Singer-songwriter	Local	Female	Indie Label
PN1	Uruguay	26-35	Alternative Rock/Indie, Progressive Rock	Local	Male	Independent
PJ1	Spain	36-45	Jazz, Free Improvisation	International	Male	Independent
PR1	Spain	36-45	Hip-Hop, Reggae, Blues, Salsa, Flamenco	International	Male	Indie Label

3.3 Processing and Analysis

Following the methodology of Qualitative Content Analysis [41], the interviews were recorded and transcribed, followed by developing an annotation scheme (i.e., coding) inductively from the transcriptions of the interviews, which we used to annotate the transcriptions accordingly.

The total duration of the recordings is 420 minutes, which transcribed correspond to 33,669 words. The annotation scheme was developed inductively from the transcriptions, where statements were used as the level annotations. Often, statements were on sentence level; yet, many sentences include two or more statements. Note that the annotation scheme was developed in English while the transcriptions were kept in their original language (i.e., Spanish). The development of the annotation scheme and the annotation of statements itself was an iterative process where we assigned a topic to a specific sentence and if it did not fit to any of the previous topics then a new one was defined. We iterate a total of 4 times and the final annotations were reviewed by a different person than the annotator to increase intercoder reliability.

The final annotation scheme includes a total of 15 overall topics that were obtained from 752 annotated text sections. In Table 3, we describe the topics with an example annotation for each topic. Note, quotes are given as translations to English, whereas the interviews were held in Spanish. Table 4 presents the number of annotations per participant per topic.

4 Results

In the first part of this section (Sections 4.1 to 4.3), we present how the artists feel affected by current music platforms. Subsequently (Sections 4.4 to 4.8), we discuss those topics that give direction on what the artists consider fair music platforms; in the paper at hand, we report those topics that were addressed by *all* participants. Table 5 provides an overview of these topics and summarizes what the participating artists deem necessary for future fair MRS.

Table 3: Details on the annotation scheme.

Topic	Description	Example of annotation
User view	Participant comments on using a music platform in the role of a music consumer.	PN1: <i>"I usually read the artist's biography and the influences of an artist."</i>
Artist view	Participant expresses opinion from the point of view of their artist role.	P2: <i>"[...] it wouldn't hurt if the systems were more random, not that obvious—if you like 'Beatles,' I recommend 'Rolling Stones.'"</i>
Lack of control	Reference to giving more control either to the artists over the music presented, or to users over what they get recommended.	P2: <i>"As an artist, I would love to have more freedom of action over my music on the platform."</i>
Diversity	Related to any aspect of diversity in the recommendation.	P3: <i>"I would expect that the music platforms promote more diverse content."</i>
Context of music	Aspects related to information and presentation apart from the music; the context that it is embedded in.	P1: <i>"There are songs that have history, [you cannot ignore that]."</i>
New music	Participant refers to artists new on a music platform or new music of existing artists.	P2: <i>"[...] it makes sense if half of the recommendations [made by the music platform are songs of] new artists."</i>
Popularity bias	Participant refers to aspects related to popularity bias of recommender systems or the music business.	P4: <i>"The problem is [that] the recommender system systematically ignores all those potential artists because it is easier to recommend [what is more popular]."</i>
Influencing behavior/ taste	Participant expresses opinion concerning music platforms' opportunity to influence users' listening behavior or music taste.	PF1: <i>"In my opinion, you can't impose some [specific] music to the users."</i>
Transparency	Refers to the need for information about how a music platform works and how its recommender system makes decisions.	PN1: <i>"If a human makes that decision if he says 'I have a small store, I am going to put it this way,' I will understand it better than if an algorithm does it."</i>
Labels' / platforms' interests	Participant refers to the interests of stakeholders such as the music platform providers or record companies.	P2: <i>"[A platform] needs to take responsibility for its recommender system [...] Obviously, they are commercially not capable or not interested."</i>
Size of artists' repertoire	Participant distinguishes between artists with more or fewer songs or albums.	PR1: <i>"If you have more songs, you have more chances to satisfy different audiences."</i>
Quotas for local music	Participant mentions regulations for local music such as quotas for local artists on music platforms.	PR1: <i>"Otherwise, you won't know what there is in your country."</i>
Gender balance	Participant talks about gender bias in the music industry or on music platforms, or how recommendations might be fair(er) from a gender perspective.	PN1: <i>"[...] the population of the world is 50% women. So it would be ridiculous if the system wouldn't recommend it."</i>
Regulation of recommendations	Participant refers to regulations or policies for music platforms or their recommender systems.	P3: <i>"[...] the question is if it should be imposed by the state [to promote local music]."</i>
Royalties distribution	Participant refers to royalties generated on music platforms or their distribution among artists.	PF1: <i>"It is absurd what [the platform] pays to the artists."</i>

Table 4: Statistics about annotations.

Topic	P4	PF1	P3	P2	P1	PN1	PJ1	PR1	PF2	Total
User view	8	1	5	8	6	5	4	2	7	46
Artist view	15	12	25	22	14	17	4	9	13	131
Lack of control	15	4	8	14	6	5	0	2	8	62
Diversity	5	2	8	6	4	2	1	3	8	39
Context of music	7	6	8	2	17	1	0	5	0	46
New music	12	6	7	8	6	14	6	10	13	82
Popularity bias	6	0	8	5	1	4	2	5	1	32
Influencing user behavior/taste	7	2	17	7	16	3	1	7	7	67
Transparency	7	0	11	13	2	6	0	2	2	43
Labels'/platforms' interests	8	4	14	17	13	5	5	11	5	82
Size of artists' repertoire	2	1	2	2	2	4	1	2	1	17
Quotas for local music	3	3	8	3	5	5	2	2	3	34
Gender balance	5	3	5	4	0	2	1	1	2	23
Regulations of recommendations	3	0	3	0	4	0	0	0	0	10
Artists' income distribution	3	1	4	9	4	4	3	7	3	38
Total	106	45	133	120	100	77	30	68	73	752

4.1 Fragmented Presentation

The participants report that they do not find adequate the way they are presented on the music platforms.

Two artists (P2 and P3) mention that their artist profiles on the music platforms show old tracks at the top because those are the most-listened items over the years. P3: *"But it is something that I have done 10 years ago. [The platform] puts the most listened tracks. [...] and you have to scroll down to reach the latest album."*

Also, the context in which artists and their music are presented may affect their public image; e.g., the artist P3 refers to a feature that serves users with an automatically generated playlist (called "radio"). The "radio" of an artist is an infinite playlist that includes tracks of this artist and also music by other artists. P3 reports that the "radio" based on his band includes music that he does not like and features artists that he distances himself from ideologically. P3: *"I see things that I do not like and that I reject ideologically. Why appears [Band X]?"* P3 explains that the band works hard on creating a certain public image—with the lyrics, the music, the art. The music platform then mixes it with something different. P3: *"I don't think the same people listen to it. That bothers me as an artist."*

P1 states that the current music platforms disconnect the music from its context. He points out that a song is inseparable from its social context. P1: *"Music—art—is a representation of people's sensitivity, it's a diary of people telling what happens."* P1: *"Listening to hip-hop from the '90s [is tied to] the slums of Los Angeles, of New York, and [to] what was happening at that time. That music comes from a place. It doesn't come out of nowhere."* P1 thinks that current music platforms do not provide much context of the music and emphasizes that including such information would enhance the experience. P1: *"[...] there are songs that have their history, their function. So the more information there is—who are the people who made that song, with whom, how, where, why they sing it—, I know it would be much richer."* Yet, P1 adds that some music may not convey any deeper message but exists for business reasons only. P1: *"[They are] made to sell more records."* He puts the example of the genre Reggaeton, for which he thinks that frequently the explicit video is the selling argument and not the music. In such a case, adding information about the context would not add much to the user experience.

4.2 Reaching an Audience

Another frequently mentioned issue is the difficulty to reach a larger audience, either because the artists are newcomers or when established artists enter a new music platform. This issue is usually known in recommenders systems as the item cold-start problem. While it has become easier than ever before to access an enormous amount of music, the artists (P2, PN1, PF1, PF2, PR1) state that it is not easy to discover less popular artists with current music platforms; it requires the user actively looking for those artists when encountering them via other sources such as magazines or interviews. PF1 feels that it is very difficult to reach a larger audience. PF1: *"If [a track] hasn't been*

listened to a number of times, it does not show up, and that means that it is not being recommended." PF2: "[...] if [the music platform] does not recommend things that [...] have rarely been listened to, then it enters a circle that never ends... it always goes... you will never be listened to more. Then you're stuck there until someone pays for you to have the promotion." P2 also participates in a less-popular band project. He affirms that music platforms make it difficult for the users to reach the band. P2: "[...] it would not hurt if the systems were more random, not that obvious—if you like 'Beatles,' I recommend 'Rolling Stones.'"

PN1 underlines that the emergence of large music platforms changed the entire music industry, making it even more difficult for new artists to reach a larger audience. PN1: "Before you could go [to a label,] with something super weird but super interesting that could catch their attention and take you on a tour [...]. [But now] it's not the music that sells. [If] you don't have followers, you don't have content, you have nothing, you're nobody, and that's why you won't appear [in the recommendations]. You have to grow in a different way, through Instagram for example, which doesn't have anything to do [with music], and the value of music gets lost."

Also, PR1 points out the difficulty of getting visibility on the music platforms if the artist is not popular. Similar to PN1, PR1 argues that it was also hard for an artist to reach visibility before the emergence of music platforms. PR1: "[...] in this business, there have always been many traps. The musician wants to make music, but it's a business. At first, you don't want to see it because you want to make music and you are happy. So for example, the old record companies used to have a monopoly before the [social] networks. You were only played on the radio if they paid for it. They said, '[This artist] has sold twenty thousand copies' but it was the record company itself that bought twenty thousand copies. [...] then everyone wanted to hear that [artist]." PR1: "On YouTube, there were companies that made it to reach the million [...] And I think that [happens] today too. You can buy visits [...]." P3 adds that artists being excluded, making it hard to reach an audience, has happened ever since, and he provides an anecdote of Bob Marley going to a radio station to force them to play his songs. P3: "They had Bob Marley and they didn't play it!" P3 argues that artists who really want to reach an audience will find a way to do it. Yet, probably not through a music platform but using other digital mediums.

In conclusion, there is no clear consensus about the actions that music platforms should take to be fair in this context. Among the mentioned alternatives, we see that some artists (e.g., P1, P2, P4, PN1) suggest that the music platforms should have a minimum quota of starting artists that are recommended to all the users alike. Others suggest that new artists should use alternative ways to reach an audience.

4.3 Transparency

Several artists (P3, PN1, P1, P2, PR1, PF2) mention that the music platforms should be more transparent. P3 states that he does not understand how exactly the music platform promotes some artists more than others; e.g., in the automatic playlists or the curated playlists. P3: "It would be nice if the platform was equitable, or fair, for everyone in that sense, because if I'm one of the largest bands in Uruguay, why I'm not in many of the playlists there? Is it the platform that doesn't want me there? Is it me doing something wrong? [...] Maybe the platform does not benefit much with what we do, so they discard us." PF2 thinks that the music platforms should be more transparent towards the artists about how their recommendation system works and what the artists have to do for being recommended more often. PF2 thinks that this is particularly important for independent artists. PF2: "[...] you are a bit naked there. You put your music on Spotify and mention in the concerts that they can listen [to your music], but you don't see any change. For example, no one explains to you that it is important that other people add your songs to their playlists so that the algorithm will recommend you."

PJ1 feels that music platforms are profitable only for some of the stakeholders. The artist wonders what the goals of the music platforms are. PJ1: "Is the goal for people to listen to music or is the goal to make money from it?" PJ1: "The people who invest in these things at the [...] powerful industry level, these people do make money from this. The others do not. [...] there is no middle class. There is an upper class and a lower class."

Regarding transparency in algorithmic decisions, the artist PN1 mentions that although both humans and algorithms may be biased, a non-ideal decision made by a human may be easier to accept than one taken by an algorithm. PN1: "[...] if a human makes that decision if he says 'I have a small store, I am going to put it this way,' I will understand it much more than if an algorithm does it."

4.4 Influencing Users' Listening Behavior

One of our interests was to learn what the artists think about the music platforms' opportunities and power to influence the users' listening behavior. For example, by tailoring the music recommendations, a music platform could balance genres or consider a gender balance among the recommended artists. In an open question, which did not mention gender, all artists came up with the issue that content by female artists is not well represented. We found that all artists agreed that the music platforms should promote content by female artists to reach a gender balance in what users consume.

While there was clear consensus to influence users' listening behavior with respect to the artists' gender—to reach a balance—, there was also a clear agreement *not* to do so concerning the music style. For the latter, they think it is better not to influence the users. P2 suggests a gender balance in the recommendations. P2: “[Platforms have] a huge responsibility in making recommendations.” Regarding gender balance, PN1 states, PN1: “[...] the population of the world is 50% women. So it would be ridiculous if the system wouldn't recommend them.” PN1 suggests a progressive change, which he thinks will prevent users to perceive it as something negative and leave the music platform. PF1 states that the system could enforce a 50% balance of male and female artists because many other factors different from gender, e.g., the music style, define whether someone will like what is recommended. Finally, PF2 considers that using quotas alone would not be enough, as there is a need for a change in education to have a bigger impact. However, she agrees that artists could be better off with quotas. PF2: “As a female artist I would like the system to recommend my music to someone that only listens to the music of male artists.”

4.5 Popularity Bias

Researchers have put a lot of effort into reducing the popularity bias and improving the recommendations for items in the long-tail of the popularity distribution (e.g., [53]). Reaching out to the artists, we wanted to explore their perspective on that issue; whether or not they feel that current systems allow the users to access items of less popular artists; and how this could be improved in the future.

P4 states that it may be easier for the music platforms to recommend what is popular because it can satisfy the majority of the users. Yet, P4 points out, P4: “The problem is [that] the recommender system systematically ignores all those potential artists [...]”. P4 thinks that some users may be happy with the recommendation of the generally popular items, whereas others may not and probably leave the music platform. The users who are more passionate about music will not see any advantage in using such recommendations. P4: “[...] it's wrong that you don't have the option to explore that long-tail. [You can't] take advantage of a recommender system if you want to [explore the long-tail].” PR1 considers that the music platforms may generate higher revenues if they recommend popular artists, and may therefore not be interested in promoting less popular content.

Although strongly advocating the promotion of diverse content, PF2 speculates that this may again lead to having users listen to widely popular music. Therefore, she claims that the music platforms should prevent it, PF2: “[...] otherwise, you will always end up listening to American music.”

All interviewees agree that it is crucial that the music platforms also recommend less popular music. They believe that the music platforms will harm the music culture if recommendations are limited to the most popular artists.

4.6 Artists' Repertoire Size

While popularity bias is a widely researched topic in the recommender systems community, and for MRS in particular, little attention is paid towards how the size of an artist's repertoire affects the probability of their songs being recommended to users. The artists' opinions are divided concerning how a music platform should reflect the differences in the repertoire sizes. While three artists think that artists with larger repertoires should be more represented in recommendations (P1, PF1, PR1), four artists do not support that idea (P2, P4, P3, PF2), and two artists were indecisive (PN1, PJ1).

P1 argues that the higher number of records leading to an increased likelihood for an artist's items being recommended reflects what happens outside the music platforms. P1: “[...] that's fine, the same thing happens in real life if someone makes 25 records, you will surely come across it at some point.” PJ1, P2, and P4, in contrast, argue that having more records should not be a reason for being

recommended more often. P2: "[...] I know so many amazing bands with only one album, it has 10 songs. And you will never get to those [being recommended]." PJ1: "This is delicate because there are big artists that have one album, or the opposite." P4: "Intuitively, I think it is unfair that an artist with more music is recommended more. [...] if an artist has 30 albums but they are all completely different from the previous ones, then it makes sense. But if there is an artist whose albums are all the same ...[it does not]."

PR1 points out that artists with more songs are probably more diverse in their repertoire. So, it is more about the diversity than the size of the repertoire. PR1: "If you have more songs you have more chances to satisfy different audiences."

P3 states that an artist profile with more songs may leave the impression that the artist is in the music business for a long time, which may be a reason to recommend the artist more. But he adds that repertoire size should not be given such importance. PF2 adds that sometimes not all tracks of an artist are available on a particular music platform, so the repertoire size on the music platform would not reflect reality.

PN1 raises the issue that it may depend on the users' goals whether the artists' repertoire size matters. If a user wants to explore a new artist, then it is not useful to recommend to this user an artist with only a few songs. For users exploring on a track basis, the artists' repertoire size is irrelevant.

4.7 Quotas for Local Music

Today's most prominent music platforms operate in several countries, more or less globally. With the widely adopted collaborative filtering approaches for recommendations ("people who like that..., also like that..."), it may happen that the music preferences of users in countries with a large number of users may influence the algorithms' outputs globally. As a result, artists that are popular in countries with a smaller user number will have fewer chances to be recommended than artists that are popular in countries with a large number of users.

While there are studies investigating the existence of local trends on global music streaming (e.g., [55]) and recommendation approaches that account for country-specific music preferences (e.g., [12]), it is not clear whether and how current music platforms consider local repertoire. Similarly, outside the music platforms, some countries define quotas of local content for radio stations while other countries do not. Yet, such quotas for radio do not apply to online music platforms, specifically not for automatic recommendations.

We asked the artists about quotas, the desirability and applicability to have quotas on online music platforms, and for automatic recommendations in particular. We also asked for potential alternative solutions to deal with local content. Three artists mention that the recommenders should have quotas for local music. Other five artists were not sure whether quotas were the right solution but emphasized that it is important that the music platforms promote local content.

PR1 agrees with quotas and indicates that PR1: "otherwise, you won't know what there is in your country." PJ1 is ambivalent. On the one hand, he believes that it is not right to base such decisions on where a person is from, on the other hand, he sees the need to promote local artists because they are at a disadvantage from the start; and with quotas, they could make up for it over time.

P1 suggests abandoning the idea of defining locality in terms of countries because national borders are not necessarily cultural borders. He proposes to define locality within a radius of a user: Artists that are in a certain radius (e.g., within a radius of 5000 km) should be given a higher weight than artists outside that radius; and within the radius, different weights again, with higher weights for the closest artists. Yet, the same artist (P1) emphasizes that quotas are a necessary measure in some countries because, otherwise, local artists would not be able to make a living solely from music. Accordingly, he suggests that quotas should also apply to automatic recommendations and proposes to use a combination of country and radius. P1: "[...] if you go to the border [between Uruguay and Brazil] [...], the Brazilian influence is greater than the Uruguayan one. So it seems to me that the radius is more representative for culture."

PN1 considers it peculiar that there is a higher chance for the user to be presented with US artists compared to local ones, even if the latter are locally famous and popular. PN1 calls for more transparency and draws an analogy to the news sector. PN1: "[...] it is like reading the news in the New York Times instead of the local newspaper. [...] you know that you're reading the New York Times or the local news. But you don't know whether it is an algorithm that makes the recommendation."

PF2 and P4 argue against quotas because this could cause users to leave the music platform if they do not like local music. Yet, both emphasize giving importance to locality. P4 suggests giving individual users the chance to choose the degree of locality they want to have. PF2 suggests

promoting local content by letting users indicate the countries they would like to receive recommendations from besides their own country. Artists could also be allowed to indicate in which countries they would like their music to be recommended. This would enable artists to reach other countries. PF2: *"As an artist, you could reach more countries if you are interested."*

P2 is unsure whether quotas are the ideal measure but emphasizes that the music platforms have responsibility for what their algorithms recommend. P2: *"[A platform provider] needs to take responsibility for its recommender system—[...] understand the situation. [...] Obviously, they are commercially not capable or not interested [...] but it would be great [...] if [they] find a way to link [...] a Yankee band with a Uruguayan band [...] make a connection that contributes. [...] [The platforms] have their share of responsibility for what they are showing or recommending. I don't know if there should be quotas [...] but [...] it would be great if there was something."*

P3 questions whether quotas should be enforced by law and suggests that music platforms take the responsibility for it. P3: *"The question is whether it should indeed be state-imposed. For things to be that way, do we have to impose it?"* P3: *"[the platform] should do what is ethically correct. [...] If I were Spotify [...] I would [promote local content] in every country."* While P3 voices concerns about whether the music platforms should be trusted in deciding what to promote more or what to promote less. Yet, for gender fairness and local content, P3 is confident that the music platforms could find the right balance.

Different from the common understanding of local quotas, PF1 suggests that it should be the opposite: Instead of having quotas for local music in smaller markets, there should be quotas in larger markets to include music from those smaller markets. In addition, PF1 points out that giving users the possibility to explore the country-specific music scene would be more beneficial for the artists. PF1: *"[...] provide the possibility to listen to what you have not listened to before. For example, 'what have I not listened to from Colombia?' 'What is underground in such a country?' If you give the local artists a voice and let them tell the story behind their music, that would be more interesting."*

4.8 New Music

New music may refer to (i) artists that are new to a user (thus, the discovery of artists) and also to (ii) a new track or album released by an artist that had already been part of the music platform.

Most of the interviewed artists agree that the artists should be in control of what tracks or albums get more recommended. In the case that they are not in control themselves, they strongly prefer a recommender system that puts more weight on their latest releases. PR1 states that every artist wants their new music to be promoted so that the world finds out they have a new release. PR1: *"[...] you do a promotion campaign to tell that you released more [content]. To tell the world, 'Hey! There is a new album!' [...] Like saying, 'Hello, I'm here.'" Also, some artists feel more identified with what they are doing now compared to music they released many years ago, which is another reason for them to prefer the promotion of the latest release. P3: "[...] it is something that I have done 10 years ago. [...] I don't know if I feel identified [with it]."*

For allowing users to discover artists that they are not yet familiar with, there is no clear consent either. While all agree—in varying degrees, though—that the music platforms should allow users to discover artists that they do not know, it remains unclear how the music platforms should do that. Most of the interviewed artists state that the user should be in control, having the opportunity to indicate that they want to discover new artists, and when they want to do so. P2: *"[...] being able to choose would be good. A button that says 'I'm open to new stuff' or 'Let me listen to what I want.' Because sometimes you want new stuff and sometimes you want something very specific."*

5 Discussion

Table 5 summarizes the concrete aspects in which the music platforms could be more beneficial for the artists. The results show that the artists' perspective on fairness relates closely to the end consumer's perspective on some aspects, whereas the interviews could also reveal that artists face problems that are not reflected in previous work on the consumer's perspective. The different roles (here, users vs item provider) may lead to varying perceptions of fairness (see [54]). For example, gender fairness is frequently also discussed from the consumer or societal perspective [50]. Our results make concrete that artists aim for *gender balance*. In contrast, the problems associated with the fragmented presentation (e.g., music presented detached from its context; artist profile may list tracks first that an artist does not identify with anymore), for example, accrue from the artist

perspective. Furthermore, the results suggest that not all identified aspects that artists consider important for a fair music platform are directly linked to algorithmic fairness. For instance, the perceived lack of control and the demand that music is presented in a way so that its context is clear are not of algorithmic nature; rather, these are system design issues. Hence, the HCI community is called to address the needs of this so-far underrepresented group.

Table 5: Aspects to improve with most of the artists in agreement.

Topic	Description
Quotas for local music	Promoting local music
Gender balance	Expectation of gender balance in recommendations
Popularity bias	Recommending items in the long tail, not only the most popular artists
Lack of control	Giving artists control concerning the tracks that are promoted; if not in control, preference for promotion of latest releases
Transparency	Transparency about how the algorithms work; why is music recommended or not
Influencing users' taste	The system should not influence the user's taste
Music in context	Music should be presented to users with information about its context

For operationalization, we can build on the strong foundations of prior research. For instance, from the interviews, we understand that artists see the need to promote new and less popular artists. While collaborative filtering is commonly used in MRS, it is an approach that is prone to popularity bias. Content-based approaches based on the advancements in music information retrieval (see [45]) could be especially apt to promote new and less popular artists.

Furthermore, as different the topics of promoting local music and ensuring gender balance in recommendations seem the basis for their operationalization exhibit similarities. First, meta-data about both, the artists' regional or cultural affiliation as well as gender information, are available for popular artists (e.g., using sources such as Wikipedia or MusicBrainz), but scarce for new and less popular artists. Thus, while existing meta-data may be used, other approaches have to be leveraged to gather missing data; this may be challenging for new and less popular artists, in particular. Second, many works investigate the diversity or coverage of computed recommendations (see [38]), little is known about how to ensure an envisaged ratio of attributes (here: region, culture, gender). In [26], we proposed a re-ranking approach to gradually achieve gender balance. Similar approaches may be used for other attributes. Targeting ratios for multiple criteria is more complex.

In addition, the finding that artists perceive their profile presentation as being fragmented and that their music is presented detached from its context, are an inspiration and rationale to consolidate and structure information from dispersed sources, so that the music presentation can be enriched with this information and put into context. For new and less popular artists, but also for new music by established artists, it will be challenging to retrieve such information. Besides challenges concerning the operationalization for information retrieval and consolidation, it is subject to future research to investigate how such contextualized information should be presented so that it (i) puts the music into context as meant by the artist and (ii) is understandable and appealing to the user. This is a great opportunity to also address the artists' demand for more control and to add interaction functionality for this stakeholder. This functionality may be as simple as an interface to insert context information or correct information that was retrieved from other sources. Providing templates for visualizing context information may add control. Further, templates and functionalities to give the artist page a personal branding may be appreciated by artists because the visual representation has long been a strong component in the music field (e.g., album artwork, stage shows, the association of genres with colors, etc.).

6 Conclusions

We reached out to music artists and conducted semi-structured interviews to understand how they are affected by current music platforms and what improvements they deem necessary so that those music platforms are *fair* from the artists' perspective. Thereby, we paid particular attention to music recommender systems that are an integral part of today's music platforms. We conclude that the participating artists' perceptions and ideas are well aligned.

The interviewed artists agree on some aspects that are needed to make the music platforms fairer: 1) The artists call for better promotion of local music; 2) they agree that gender balance in the recommendations is indisputably expected; 3) the artists voice that music items in the long tail of the popularity distribution (not only the most popular artists) have to be included in the recommendations shown to users; 4) the participants advocate giving control to the artists over the tracks that are promoted or higher weighted in recommendations (if they are not directly in control, then they generally favor the promotion of their latest releases); 5) they request transparency about how the algorithms work, to understand why their music is recommended or not; 6) the artists consider a system that influences a user's taste (or attempts to do so) an undesired misuse; and 7) they would appreciate if the music platforms would be enriched with information that puts the music into context.

Besides the consensus on many topics, there is no clear direction for others: 1) There is no agreement concerning quotas for local content; 2) no clear majority whether the size of the artists' repertoire should be reflected in recommendations; 3) while the artists seem to agree that new artists should be given space on the music platforms, there is no agreement on how to operationalize this; and there is 4) no clear agreement whether a music platform should promote the discovery of artists previously unknown by a user. Overall, while there is a prevailing belief that music platforms would open up the long-tail to users and encourage them to consume more of those items, the interviews suggest that the long-tail items and artists remain obscure.

A limitation of our work relates to the chosen context. We deliberately drew the sample from a big and coherent music market (Spanish-speaking music market), which is geographically dispersed (Europe, North America, and South America), and where the artists are diverse in styles, career, and popularity. While we could find shared ideas and opinions across the sample which is diverse in many aspects, the sample's perception could still be aligned from a cultural perspective (e.g., Uruguay and Spain are similar in most of Hofstede's cultural dimensions [32]). Yet, this defined setting (artists from Spanish-speaking countries) is a good opportunity for upcoming research to draw comparisons to other defined groups of artists. Future research may investigate the identified aspects in different contexts and more depth. Further, while we catered for some aspects of diversity in our sample of music artists, reaching out to a larger sample will allow for even more aspects of diversity (e.g., including artist of non-binary gender, considering solo artists, mixed-gender bands, mono-gender bands, various ethnic groups, artists dedicated to niche music). Although we reached thematic saturation, reaching out to a wider set of diverse artists may reveal additional topics or different viewpoints.

These caveats notwithstanding, our work gives direction towards relevant topics for fairness on music platforms and their integrated MRS. The findings indicate pathways towards fairer music platforms, whereas the concrete operationalization is subject to further research.

Naturally, the music domain has its specificities. Yet, our study results may inspire studies in other domains. First, our findings indicate that different roles may come with different fairness perceptions and requirements. Hence, we encourage studies in other domains where an item provider typically has several items and where a person is the face to the public for themselves, an organization, or a brand (e.g., sports, many technology companies). Second, our findings suggest that perceived unfairness also is tied to a lack of control. Likely, control may also be given with more opportunities for interaction.

Acknowledgments

This research was partially supported by Kakao Corp.

References

- [1] Aguiar, L.: Let the music play? Free streaming and its effects on digital music consumption. *Information Economics and Policy* **41**, 1–14 (2017). <https://doi.org/10.1016/j.infoecopol.2017.06.002>
- [2] Aguiar, L., Waldfogel, J.: Platforms, promotion, and product discovery: Evidence from spotify playlists. Tech. rep., National Bureau of Economic Research (2018)

- [3] Akimchuk, D., Clerico, T., Turnbull, D.: Evaluating recommender system algorithms for generating local music playlists (2019), <https://arxiv.org/abs/1907.08687>
- [4] Andersen, K., Knees, P.: Conversations with expert users in music retrieval and research challenges for creative mir. In: Proc. of the 17th International Society for Music Information Retrieval Conference. pp. 122–128. ISMIR '16 (2016)
- [5] Anderson, A., Maystre, L., Anderson, I., Mehrotra, R., Lalmas, M.: Algorithmic effects on the diversity of consumption on spotify. In: Proc. of The Web Conference 2020. pp. 2155–2165. WWW 2020 (2020)
- [6] Anderson, C.: The long tail. Wired (Jan 2004), <https://www.wired.com/2004/10/tail/>
- [7] Anderson, C.: *The Long Tail: Why the Future of Business is Selling Less of More*. Hyperion, New York, NY, USA (2006)
- [8] Baeza-Yates, R.: Data and algorithmic bias in the web. In: Proceedings of the 8th ACM Conference on Web Science. WebSci '16, ACM, New York, NY, USA (2016). <https://doi.org/10.1145/2908131.2908135>
- [9] Bauer, C.: Allowing for equal opportunities for artists in music recommendation: a position paper. In: Proc. of the 1st Workshop on Designing Human-Centric Music Information Research Systems. pp. 16–18. wsHCMIR '19 (2019)
- [10] Bauer, C.: Report on the ISMIR 2020 special session: How do we help artists? ACM SIGIR Forum **54**(2), 1–7 (2020), <http://sigir.org/wp-content/uploads/2020/12/p15.pdf>
- [11] Bauer, C., Kholodylo, M., Strauss, C.: Music recommender systems challenges and opportunities for non-superstar artists. In: Proc. of the 30th Bled eConference. pp. 21–32 (2017)
- [12] Bauer, C., Schedl, M.: Global and country-specific mainstreamness measures: Definitions, analysis, and usage for improving personalized music recommendation systems. PLOS ONE **14**(6), 1–36 (2019). <https://doi.org/10.1371/journal.pone.0217389>
- [13] Baym, N.K.: *Playing to the crowd: Musicians, audiences, and the intimate work of connection*, vol. 14. NYU Press, New York, NY, USA (2018)
- [14] Biega, A.J., Gummadi, K.P., Weikum, G.: Equity of attention: Amortizing individual fairness in rankings. In: The 41st International ACM Conference on Research & Development in Information Retrieval. pp. 405–414. SIGIR '18 (2018)
- [15] Binns, R.: On the apparent conflict between individual and group fairness. In: Proc. of the 2020 Conference on Fairness, Accountability, and Transparency. pp. 514–524. FAT* '20 (2020). <https://doi.org/10.1145/3351095.3372864>
- [16] Celma, O.: *Music recommendation and discovery: The Long Tail, Long Fail, and Long Play in the Digital Music Space*. Springer, Berlin, Heidelberg, Germany (2010). <https://doi.org/10.1007/978-3-642-13287-2>
- [17] Coelho, M.P., Mendes, J.Z.: Digital music and the “death of the long tail”. *Journal of Business Research* **101**, 454–460 (2019)
- [18] Cramer, H., Garcia-Gathright, J., Reddy, S., Springer, A., Takeo Bouyer, R.: Translation, tracks & data: An algorithmic bias effort in practice. In: Extended Abstracts of the 2019 Conference on Human Factors in Computing Systems. pp. 1–8. CHI EA '19 (2019). <https://doi.org/10.1145/3290607.3299057>
- [19] Cramer, H., Garcia-Gathright, J., Springer, A., Reddy, S.: Assessing and addressing algorithmic bias in practice. *Interactions* **25**(6), 58–63 (2018)
- [20] Creswell, J.W., Poth, C.N.: *Qualitative inquiry and research design: Choosing among five approaches*. Sage Publications, Thousand Oaks, CA, USA (2016)
- [21] Dwork, C., Hardt, M., Pitassi, T., Reingold, O., Zemel, R.: Fairness through awareness. In: Proc. of the 3rd Innovations in Theoretical Computer Science Conference. pp. 214–226. ITCS '12 (2012). <https://doi.org/10.1145/2090236.2090255>

- [22] Farnadi, G., Kouki, P., Thompson, S.K., Srinivasan, S., Getoor, L.: A fairness-aware hybrid recommender system. arXiv preprint arXiv:1809.09030 (2018), <https://arxiv.org/abs/1809.09030>
- [23] Ferraro, A., Bogdanov, D., Serra, X., Yoon, J.: Artist and style exposure bias in collaborative filtering based music recommendations. In: Proc. of the 1st Workshop on Designing Human-Centric Music Information Research Systems. pp. 8–10. wsHCMIR '19 (2019)
- [24] Ferraro, A., Jannach, D., Serra, X.: Exploring longitudinal effects of session-based recommendations. In: 14th ACM Conference on Recommender Systems. pp. 474–479. RecSys '20 (2020). <https://doi.org/10.1145/3383313.3412213>
- [25] Ferraro, A., Jeon, J.H., Kim, B., Serra, X., Bogdanov, D.: Artist biases in collaborative filtering for music recommendation. In: Machine Learning for Media Discovery Workshop at International Conference on Machine Learning (ICML) (2020)
- [26] Ferraro, A., Serra, X., Bauer, C.: Break the loop: Gender imbalance in music recommenders. In: 6th ACM SIGIR Conference on Human Information Interaction and Retrieval. pp. 249–254. CHIIR '21, ACM, New York, NY, USA (2021). <https://doi.org/10.1145/3406522.3446033>
- [27] Feuerriegel, S., Dolata, M., Schwabe, G.: Fair ai. *Business & Information Systems Engineering* **62**(4), 379–384 (2020). <https://doi.org/10.1007/s12599-020-00650-3>
- [28] Fleder, D., Hosanagar, K.: Blockbuster culture's next rise or fall: The impact of recommender systems on sales diversity. *Management science* **55**(5), 697–712 (2009)
- [29] Guest, G., Bunce, A., Johnson, L.: How many interviews are enough?: An experiment with data saturation and variability. *Field Methods* **18**(1), 59–82 (2006). <https://doi.org/10.1177/1525822X05279903>
- [30] Harrison, G., Hanson, J., Jacinto, C., Ramirez, J., Ur, B.: An empirical study on the perceived fairness of realistic, imperfect machine learning models. In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency. pp. 392–402. FAT* '20, ACM, New York, NY, USA (2020). <https://doi.org/10.1145/3351095.3372831>
- [31] Helberger, N., Araujo, T., de Vreese, C.H.: Who is the fairest of them all? public attitudes and expectations regarding automated decision-making. *Computer Law & Security Review* **39**, 105456 (2020). <https://doi.org/https://doi.org/10.1016/j.clsr.2020.105456>
- [32] Hofstede, G., Hofstede, G.J., Minkov, M.: *Cultures and organizations: Software of the mind*, vol. 3rd, revised. McGraw-Hill, New York, NY, USA (2010)
- [33] Holstein, K., Wortman Vaughan, J., Daumé, H., Dudik, M., Wallach, H.: Improving fairness in machine learning systems: What do industry practitioners need? In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. CHI '19, ACM, New York, NY, USA (2019). <https://doi.org/10.1145/3290605.3300830>
- [34] Holzapfel, A., Sturm, B., Coeckelbergh, M.: Ethical dimensions of music information retrieval technology. *Transactions of the International Society for Music Information Retrieval* **1**(1), 44–55 (2018)
- [35] Hutchinson, B., Mitchell, M.: 50 years of test (un)fairness: Lessons for machine learning. In: Proc. of the Conference on Fairness, Accountability, and Transparency. pp. 49–58. FAT* '19 (2019). <https://doi.org/10.1145/3287560.3287600>
- [36] Jannach, D., Bauer, C.: Escaping the mcnamara fallacy: Toward more impactful recommender systems research. *AI Magazine* **41**(4), 79–95 (2020). <https://doi.org/10.1609/aimag.v41i4.5312>
- [37] Kowald, D., Müllner, P., Zangerle, E., Bauer, C., Schedl, M., Lex, E.: Support the underground: Characteristics of beyond-mainstream music listeners. *EPJ Data Science* **10**(1) (2021). <https://doi.org/10.1140/epjds/s13688-021-00268-9>
- [38] Kunaver, M., Požrl, T.: Diversity in recommender systems – a survey. *Knowledge-Based Systems* **123**, 154–162 (2017). <https://doi.org/10.1016/j.knosys.2017.02.009>

- [39] Madaio, M.A., Stark, L., Wortman Vaughan, J., Wallach, H.: Co-designing checklists to understand organizational challenges and opportunities around fairness in ai. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. CHI '20, ACM, New York, NY, USA (2020). <https://doi.org/10.1145/3313831.3376445>
- [40] Marlin, B., Zemel, R.S., Roweis, S., Slaney, M.: Collaborative filtering and the missing at random assumption. arXiv preprint arXiv:1206.5267 (2012), <https://arxiv.org/abs/1206.5267>
- [41] Mayring, P.: Qualitative content analysis. In: A companion to qualitative research, chap. 5.12, pp. 159–176. SAGE, London, United Kingdom (2004)
- [42] Mehrotra, R., McInerney, J., Bouchard, H., Lalmas, M., Diaz, F.: Towards a fair marketplace: Counterfactual evaluation of the trade-off between relevance, fairness & satisfaction in recommendation systems. In: Proc. of the 27th ACM International Conference on Information and Knowledge Management. pp. 2243–2251. CIKM '18 (2018). <https://doi.org/10.1145/3269206.3272027>
- [43] Milano, S., Taddeo, M., Floridi, L.: Recommender systems and their ethical challenges. AI & Society (2020). <https://doi.org/10.1007/s00146-020-00950-y>
- [44] Morse, J.M.: Designing funded qualitative research. In: Handbook of qualitative research, pp. 220–235. Sage Publications, Thousand Oaks, CA, USA (1994)
- [45] Murthy, Y.V.S., Koolagudi, S.G.: Content-based music information retrieval (cb-mir) and its applications toward the music industry: A review. ACM Computing Survey **51**(3) (2018). <https://doi.org/10.1145/3177849>
- [46] Rosen, S.: The economics of superstars. The American Economic Review **71**(5), 845–858 (1981), <http://www.jstor.org/stable/1803469>
- [47] Sapiezynski, P., Zeng, W., E Robertson, R., Mislove, A., Wilson, C.: Quantifying the impact of user attention on fair group representation in ranked lists. In: Proc. of The 2019 World Wide Web Conference. pp. 553–562. WWW '19 (2019). <https://doi.org/10.1145/3308560.3317595>
- [48] Selbst, A.D., Boyd, D., Friedler, S.A., Venkatasubramanian, S., Vertesi, J.: Fairness and abstraction in sociotechnical systems. In: Proc. of the Conference on Fairness, Accountability, and Transparency. pp. 59–68. FAT* '19 (2019). <https://doi.org/10.1145/3287560.3287598>
- [49] Singh, A., Joachims, T.: Fairness of exposure in rankings. In: Proc. of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. pp. 2219–2228. SIGKDD '18 (2018). <https://doi.org/10.1145/3219819.3220088>
- [50] Sonboli, N., Smith, J.J., Berenfus, F.C., Burke, R., Fiesler, C.: Fairness and transparency in recommendation: The users' perspective. arXiv preprint arXiv:2103.08786 (2021), <https://arxiv.org/abs/2103.08786>
- [51] Srivastava, M., Heidari, H., Krause, A.: Mathematical notions vs. human perception of fairness: A descriptive approach to fairness for machine learning. pp. 2459–2468. KDD '19, ACM, New York, NY, USA (2019). <https://doi.org/10.1145/3292500.3330664>
- [52] Turnbull, D., Waldner, L.: Local music event recommendation with long tail artists (2018), <https://arxiv.org/abs/1809.02277>
- [53] Vall, A., Quadrana, M., Schedl, M., Widmer, G.: Order, context and popularity bias in next-song recommendations. International Journal of Multimedia Information Retrieval **8**(2), 101–113 (2019). <https://doi.org/10.1007/s13735-019-00169-8>
- [54] Wang, R., Harper, F.M., Zhu, H.: Factors influencing perceived fairness in algorithmic decision-making: Algorithm outcomes, development procedures, and individual differences. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. pp. 1–14. CHI '20, ACM, New York, NY, USA (2020). <https://doi.org/10.1145/3313831.3376813>
- [55] Way, S.F., Garcia-Gathright, J., Cramer, H.: Local trends in global music streaming. In: Proc. of the International AAAI Conference on Web and Social Media. vol. 14, pp. 705–714 (2020)

- [56] Woodruff, A., Fox, S.E., Rousso-Schindler, S., Warsaw, J.: A qualitative exploration of perceptions of algorithmic fairness. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. pp. 1–14. CHI '18, ACM, New York, NY, USA (2018). <https://doi.org/10.1145/3173574.3174230>
- [57] Yao, S., Huang, B.: Beyond parity: Fairness objectives for collaborative filtering. In: *Proc. of the 31st International Conference on Neural Information Processing Systems*. pp. 2925–2934. NIPS'17 (2017). <https://doi.org/10.5555/3294996.3295052>