‘Don’t mess with my algorithm’: Exploring the relationship between listeners and automated curation and recommendation on music streaming services

by Sophie Freeman, Martin Gibbs, and Bjørn Nansen

Abstract
Given access to huge online collections of music on streaming platforms such as Spotify or Apple Music, users have become increasingly reliant on algorithmic recommender systems and automated curation and discovery features to find and curate music. Based on participant observation and semi-structured interviews with 15 active users of music streaming services, this article critically examines the user experience of music recommendation and streaming, seeking to understand how listeners interact with and experience these systems, and asking how recommendation and curation features define their use in a new and changing landscape of music consumption and discovery. This paper argues that through daily interactions with algorithmic features and curation, listeners build complex socio-technical relationships with these algorithmic systems, involving human-like factors such as trust, betrayal and intimacy. This article is significant as it positions music recommender systems as active agents in shaping music listening habits and the individual tastes of users.

Contents
Introduction
Methods
Findings and discussion
Conclusion

Introduction
Music streaming services offer on-demand delivery of music audio that front-end users access via freemium or ad-supported subscriptions (Sinnreich, 2016). In major shifts from the traditional models of the music industry, music streaming services no longer require users to download or pay for individual albums and songs (Wikström, 2020). The shifting technology has brought about immense changes to the way that music
is found, stored and listened to. For instance, popular streaming platform Spotify reports 381 million active monthly users (as of November, 2021) and a powerful music recommendation system that curates and personalises playlists based on user data and ‘taste profiles’ (Spotify, n.d.a). Access to huge online collections of music means users become increasingly reliant on algorithmic recommender systems and automated discovery features to find and curate music (Barna, 2017). Through both human and algorithmic curation, Spotify’s popular in-house curated playlists continue to grow in popularity and number and have become integral to artist strategies for promotion (Morgan, 2020). With a huge music and podcast catalogue, data analytics and recommendation capacity, Spotify is undoubtedly shifting consumption and distribution in the music industry. With access to seemingly infinite (Burkart and McCourt, 2006) catalogues of music (Spotify boasts more than 70 million songs), users increasingly interact with or rely on algorithmic filtering, curation and recommendation to find music (Spotify, n.d.a).

As in many other algorithmic, and proprietary contexts, the exact logic of the Spotify recommendation engine is a commercial secret (Bonini and Gandini, 2019). Despite this, some insights can be gleaned through industry conferences, music reporting (Pelly, 2019), industry blogs and from the company themselves (Spotify, 2020). Major music streaming services, such as Spotify and Apple Music, use data-driven and automated curation, alongside editorial selection to recommend music, personalise playlists and ultimately decide which artists are made visible to listeners (Bonini and Gandini, 2019). Morris (2020) points to the “platform effects” on music streaming services, whereby competing and conflicting agendas from users, content vendors, and service providers meet, creating dynamic relationships and practices through their various attempts to optimise music. Spotify’s explanations of how they recommend content hints at these commercial agendas, which may encroach on other reasons to optimise music:

“Our personalised recommendations are tailored to your unique taste, taking into account a variety of factors, such as what you’re listening to and when, the listening habits of people who have similar taste in music and podcasts, and the expertise of our music and podcast specialists. In some cases, commercial considerations may influence our recommendations, but listener satisfaction is our priority and we only ever recommend content that we think you’ll want to hear.”

(Spotify, n.d.b)

This explanation from Spotify offers some insight into how content is recommended to listeners, but it also highlights a tension between algorithmic and commercial logics of the platform. Within this same context, Prey (2020) notes that music streaming services serve more than one market: serving the music, advertising and financial markets. For Prey (2020) these competing agendas are crystallised in playlists on Spotify through both the algorithmic and editorial logics used to create them, but also through underlying financial pressures specific to these platforms embedded within multiple markets. With music streaming services emerging as powerful tastemakers, users increasingly rely on automated features to find and curate music for them, and music streaming platforms inherently gain more power over what content becomes popular (Arielli, 2018). This is problematic as the Spotify system is the market leader in the global music streaming market, and has affected real changes in music consumption, discovery and industry behaviour (Eriksson, et al., 2019). In keeping with an emerging algorithmic culture, spanning many areas of human activities, tasks that were traditionally performed by humans, such as discovery, recommendation and curation, are increasingly being relegated to algorithmic filtering, machine learning and AI (Gillespie, 2017). Gatekeepers in the music industries are not new. Radio DJs, music labels, record stores, and Billboard charts have variously asserted power and dictated who ‘makes it’ (Brewster and Broughton, 2014). However, recommendation and discovery have never occurred in such a personalised and branded way as they do on commercial music streaming services (Eriksson, et al., 2019). In the music streaming context, processes of discovery, curation and recommendation are increasingly automated and personalised, with music being optimised for accuracy according to both algorithmic and commercial logics (Hracs and Webster, 2021). This raises compelling questions around user agency in choosing what to listen to, and when.
Algorithmically powered features and recommendations offer new possibilities for music discovery and for distribution, at the same time as prescribing certain discovery and listening experiences (McKelvey and Hunt, 2019). Conceptualising the interaction between user and algorithmic recommender system as a relationship, this research situates this particular human-machine relationship within broader contexts of algorithmic culture. Following a similar call from Jansson (2021), to pay attention to the ways streaming platforms are embedded in social and power relations, we unpack some complexities of this particular algorithmic system by exploring the user experience of music streaming. This research begins to unpack the complex, and dynamic, socio-technical relationship that forms between user and recommender system, asking how they understand one another, and how this understanding shapes their use [1].

**Recommendation & curation on Spotify**

Streaming services shift listeners’ experiences from private, digital collections on individual devices to public profiles, via ‘access’ to a cloud-based catalogue and to personalisation on an unprecedented scale. In order to make personalised recommendations, Spotify employs a unique combination of machine and deep learning models, coupled with human curation, to process and interpret musical attributes, and behavioural (listening) user data in order to offer recommendations and personalisation (Goldschmitt and Seaver, 2019). Spotify uses a hybrid method for recommendation: a combination of techniques such as collaborative filtering, contextual and content-based recommendation (Celma, 2010; Goldschmitt and Seaver, 2019). Data pertaining to how the music sounds (acoustic), who listens to it (user), when and where they listen to it (contextual), what do they do with it, i.e., playlisting (behavioural) data is combined with natural language processing of what people say about the music on thousands of sites (semantic music knowledge) (Spotify, 2020). Taken together this combines to form a listener’s ‘taste profile’ (Sónar +D, 2016) and is reflected back to users in the form of recommendations.

This combination of acoustic, user, contextual, behavioural and semantic music knowledge data allows for tailored personalisation and algorithmic curation and recommendation. For instance, Spotify offers a number of personalised playlists under their Made For You categories, such as Your 2021 Wrapped, Daily Mixes, Release Radar and more. Spotify’s Discover Weekly is a 30-song hyper-personalised playlist, algorithmically curated each week which Spotify describes as, “Songs we think you’ll love” (Spotify Support, n.d.).

*Daily Mixes* are six continuous playlists that create a mix of liked artists and genres. *Release Radar* algorithmically collates the new music from a users top artists (liked, following, listened to). These are just some of the algorithmic personalisation and recommendation features on the Spotify platform, and curation across the platform is often a hybrid of human and algorithmic. In their study into curation on music streaming platforms, Bonini and Gandini (2019) find that it is through a combination of algorithmic filtering and human curation that songs are selected for playlists. Human curators read listener data and use this to curate, populate or moderate playlists. This reliance on data and analytics allows human curators to follow the data. Music streaming platforms, through their branded offerings and features powered by algorithmic personalisation and curation, set the boundaries and possibilities for music listening within their software and interfaces. More broadly, with their powerful analytic and algorithmic capabilities music streaming services have the potential to enable, inform and prescribe the cultural practices and experiences of and with music (Morris and Powers, 2015). ‘Taste’ and ‘mood’ are shifted from subjective human traits and experiences to commercial and algorithmic data points for recommendation and advertising (Prey, 2019). Within this infrastructure, previously subjective human processes of curation and discovery are increasingly mediated by algorithmic systems.

**Control and transparency**

Spotify offers what Eriksson, *et al.* (2019) call a ‘commodified experience’, in their extended critique and analysis of the ways that the commercial service may be affecting cultural and music practices. Journalist
Don't mess with my algorithm: Exploring the relationship between listeners and automated curation and recommendation on music streaming services

Liz Pelly (2019) has described Spotify as a ‘big mood machine’ whereby the company uses emotional surveillance for profit, or as ‘streambait’ in which the service homogenises genres to gain and maintain attention. Similarly, other public critics voice fears that through its extensive personalisation and recommendation features, Spotify may trap users in a feedback loop, listening to the same or homogenous music, resulting in what McCafferey (2016) has described as a ‘taste tautology’. Others suggest that computational systems such as Spotify cannot capture or emulate the inherently human and subjective experiences of music consumption, discovery and taste (Turner, 2016). By speaking with members of the Australian music industry, Ben Morgan (2020) elucidates some of the challenges the music industry face in trying to navigate the commercially controlled promotional space of the in-house curated Spotify playlist. Taken together, this short list demonstrates just some of the challenges faced by numerous actors and user groups in relation to music streaming, automated curation and algorithmic recommendation.

Jeremy Wade Morris (2015) describes music recommendation algorithms as info-mediaries, pointing to the power of algorithmic curation to shape not just what content is accessed, but also how it is received. Morris (2015) posits that when “we turn culture into data” it affects content, consumption and taste. Similarly, Gillespie (2017) argues that music recommender systems are both predictive and prescriptive, creating recursive loops of content. Echoing the concerns of Liz Pelly (2018), Eriksson, et al. (2019) and other critics of music streaming services, for Nowak and Whelan (2016) on music streaming services “consumption practices are set by corporate monopolies”. Further exploring the effects of these corporate monopolies, Robert Prey (2018) applies Gilbert Simondon’s theory of individuation to ‘ways of seeing’ the individual on music streaming services. Prey (2019) argues that these ways of seeing are heavily influenced by the consumer categories that are defined and demanded by advertisers. Behavioural listening data is aggregated and leveraged for music recommendation (as well as for third party advertising). Taken together, the automated curation and recommendation features on music streaming services are programmed with the commercial imperative to keep listeners listening (Seaver, 2017).

The increasing dependence on algorithms and automated curation to personalise experiences and discover music on behalf of listeners presents not only a deviation from the music industry and technologies of the past (characterised by traditional industry gatekeepers and personal collections, both digital and analogue), but presents a complex socio-technical dynamic, whereby algorithms in the context of music streaming services can be seen as active agents shaping the individual tastes, discoveries and listening experiences of users. In order to explore this dynamic, this research asks: How do listeners experience music recommender systems? How do users understand and conceptualise algorithmic systems, and how does this understanding shape their use?

Algorithmic relations in music listening

This article takes cues from Taina Bucher’s (2016) work on the algorithmic imaginary, where she asks how people understand and ‘know’ algorithmic systems. Similarly, this article builds on the work of Siles, et al. (2020) who look to the Costa Rican context, exploring folk theories of algorithmic recommendations on Spotify as a lens through which to view culturally specific data assemblages. Using music streaming and recommender site Last.fm as an example, Lury and Day (2019) discuss the logic of collaborative filtering and the ways that personalisation through recommendation constrains and controls use, and indeed user identities. Pointing to Gillespie’s (2014) ‘entanglement of algorithms with practice’, Karakayali, et al. (2018) posit that algorithms are technologies of the self, as ‘intimate experts’ or at least as ‘companions’.

A growing number of studies focus on playlists and playlisting behaviours on streaming platforms. For instance, Hagen (2015) used diary methods and interviews to understand user playlisting practices. Siles, et al. (2019) look to the way that playlists cultivate moods and emotions. Through group interviews with music streamers in Norway, Kjus (2016) looks to users’ sense of music discovery through streaming. Looking more towards industry, Maasø (2018) studies the ‘eventisation of Spotify’ in Sweden by studying how artists were streamed before, during and after live music festival performances. Morgan (2020) looks to attitudes and techniques used by artists and managers to leverage in house curated (and owned) playlists for promotion on music streaming platforms.
Although there are increasingly more ‘user-centric’ evaluations of music recommender systems in the fields of music information retrieval (MIR) and human-computer interaction (HCI) these tend toward quantitative methods, implicit evaluations and off-line experiments, and tend away from critical evaluations of commercial recommender systems (Schedl, et al., 2018). Media studies approaches look critically at algorithmic systems, often perpetuating the inside/outside dichotomy inherent in black box conceptualisations (Pasquale, 2015). These critical studies of platform and system are certainly valuable, but do not always address the user experience, and the daily interactions between listeners and algorithmic features. Likewise, research into the relationship between users and algorithms is increasing, with reference to folk theories and user knowledge of algorithms, though these have focused on other social media such as Facebook (Bucher, 2016; DeVito, et al., 2018; Eslami, et al., 2016), rather than music streaming contexts. Despite growing literature into music recommendation and streaming across platform studies, critical algorithm studies and MIR little is known about the user experience of Spotify, or indeed the socio-technical practices emerging around and through these platforms and algorithmic systems. By addressing the user experience of listeners through interviews and participant observation, this article aims to fill this gap by studying the new and complex relationships form between users and recommender systems. These relationships demonstrate that there is space for tensions between human subjectivity and algorithms given human-like agency over music listening and discovery.

Methods

Music curation on contemporary music platforms presents challenges to academic research, which Seaver (2017) attributes to lack of access to the field of research, and Bucher (2016) suggests is due to both epistemological and methodological constraints that form when conceiving of algorithmic systems as black boxes. This research is concerned with the ways that algorithmic recommendation and curation shape music taste and socio-technical practices, and as such focuses on the individual experience of listeners (cf., Jansson, 2021). Semi-structured interviews and participant observation were chosen to extract rich qualitative data about complex systems and individual experiences (Hennink, et al., 2020). The interviews were organised and analysed thematically, using the six stages of thematic analysis proposed by Braun and Clarke (2012).

In conducting semi-structured interviews with participants (n=15) from the ages of 18–35, we were able to focus on users of online music streaming platform that have lived through their advent and rise in popularity. Questions focused on music recommendation and discovery across platforms, prompting interviewees to share their experiences of interacting with music recommender systems. Questions asked about daily habits and interactions with the music streaming services, as well as a particular focus on recommendation and discovery features. Questions were designed to get at perceptions and understandings of how the system was working. Interviews were conducted between late 2019 and early 2020.

Drawing from Robards and Lincoln (2017) scroll-back method, participants were invited to share their algorithmically curated and personalised playlists on their own devices, such as their Discover Weekly on Spotify or For You on Apple Music and to scroll through and discuss their reactions to the curated offerings. This allowed for participants “to fill in blanks and provide context [and] the deeper meanings” of their profiles and experiences on the app [4]. Participants displayed high interest in music and high interest in technology and were recruited as ‘daily users of a streaming platform’. Many participants played instruments, and eight participants had a professional or hobbyist link to the music industry (DJ or radio).

In sharing their use of music streaming platforms, participants offered rich narratives pertaining to their music and discovery experiences both on and off the music streaming platform. Spotify was the dominant platform among participants (n=13), followed by Apple Music which was used exclusively by two participants. As such, most of the focus of this paper is on Spotify, but other music platforms were
described across all the interviews (YouTube, SoundCloud, Mixcloud, Bandcamp). As an inherent condition of use of music streaming services listeners must engage with algorithmic recommendation and curation features on some level, which can alter or shape their experiences of music consumption and discovery. Listeners often formed a relationship with the system, describing humanlike feelings of intimacy, betrayal and trust. These daily interactions between listener and music streaming service were often reminiscent of a human relationship; one that could become fraught with issues of reliance, blame, misunderstanding and so on.

In this way, use of these systems was seen to be a negotiation, and dynamic in nature. Depending on context, mood or intention, users conceived of the recommender systems either as tools or as agents. This relationship was fluid and dynamic, both between users in this sample, and within users themselves.

Findings here are structured around the three themes of user-algorithmic relations: Alterity, Theories, and Tactics. Within each theme, participants gave some understanding of their daily habits and experiences, as well as their understanding of and epistemological relationship to the recommender system.

Findings and discussion

**Theme 1: Alterity**

The emergence of this theme demonstrated that users were conceiving of a kind of relationship between them and the system, often by giving the recommender system and platform an identity or having an imagined conversation with it. In line with Ihde’s (1990) notion of alterity relations between humans and technology, users conceptualised the recommender system as other to them. Ihde (1990) proposes that alterity relations between human and machine involve a ‘quasi-otherness’, whereby the technology and human are conceived of as separate, yet relational entities. This was a useful lens through which to analyse the relationship between participants and the algorithmic music systems they were interacting with daily. This sense of a “quasi-otherness” was described in interviews through the relationship that forms between user and system, which could be compared to a human relationship; a relationship involving reliance, trust, understanding, betrayal and other human emotions.

1.1. Characterisation + anthropomorphic descriptions

Throughout all the interviews, and indeed within individual conversations themselves, participants used varying language and metaphors to describe the recommender system they interacted with daily. In discussing the way that the system worked, many explanations were attributed to ‘the algorithm’, ‘algorithms’, simply to ‘Spotify’ and ‘they’, and also to ‘my algorithm’. In the latter, users pointed to a sense of ownership, especially when expressing the dangers of other people (friends, partners, kids) “messing with their algorithms” by listening to music while logged into their account. Using the phrase “my algorithm” also spoke to an awareness of, and often an appreciation for the personalisation that is afforded through the algorithmic recommendations made according to a user’s taste profile. The breadth of language used speaks to the range and levels of understanding about how the recommender systems work, as well as some knowledge of machine learning and recommendation in general (i.e., there is not one Spotify algorithm, but many). This level of understanding is not necessary for use of these streaming platforms, though many users were increasingly aware of the presence of algorithms.

When discussing, “their algorithm” listeners were pointing to an intimacy, or at the least a relationship that forms between them and what (or who) they perceive to be giving them recommendations. Marc described a relationship with “his Spotify” that was akin to a sibling:

“I would say that it’s almost at times like an eldest sibling that knows you better than you know yourself and therefore can
recommend you things that are maybe from like from left field.
And sometimes you kinda hear those things and you’re like, ‘I
can see where you’re coming from’.” Marc.

Others described the music recommender as a ‘personal DJ’ (Sarah), and another said that in some ways Spotify knew them better than their partner did. Because of their personalised nature, recommendations could in this way can be uncannily accurate. In the above quote, Marc described the system knowing him better than he knows himself. From this perceived level of understanding about their music taste, many users would lean on or come to allow the platform to choose music for them in many situations. Many offered conceptualisations of “someone” choosing or taking the reins for them. This was especially true when participants were engaged in other activities or were in a mood for background listening. In those particular contexts or moods, listeners were often happy to relinquish control to an agent or figure who “knew them”, and thus could be trusted to curate the music on their behalf.

1.2. Understanding & intimacy

Many participants expressed a sense of intimacy between them and the recommender system which was perceived through the understanding they attributed to the system; intimacy was often conflated with a sense of understanding, by their recommender, of them, the listener, and their taste. Ryo explained:

“I have been on a home screen previously and thought to myself, ‘Oh, this is a reasonable interpretation from someone who hasn’t met or doesn’t know me as a person’. And then if I would listen to other genres, like I also really like rap, but nowhere near to the same level as those other three [main] categories. And then it felt like Spotify was like ‘you’d been listening to rap again haven't you?!’. And that would be like Daily Mix 4...” Ryo.

In this quote Ryo is highlighting that Spotify over time built a ‘reasonable’ interpretation of him and his main music tastes (across the three broad genres he identified), but that occasionally when deviating from these genres Spotify was ‘watching’ and even ‘commented’ on this return to a forgotten genre through the new Daily Mix made up of rap music. This speaks to an understanding that the system is collecting data and following one’s listening habits. This wasn’t always a problem, in fact Layla relished the shared history between her and Spotify, suggesting that she couldn’t fathom moving to another platform and losing that understanding. Like her, others described the labour put into the platform through liked songs, saved albums, playlists and collaborations, and years of accumulated behavioural data. The listening history and personal archives in the form of playlists were described by Arianna, Boz and Taylor as ‘irreplaceable’. Layla described the archive of saved music as a narrative of her life, not just a log of behavioural data.

1.3. Reliance

In continuing the metaphor of a human relationship, many listeners felt that they relied on their chosen streaming platform in a couple of different ways. Boz described his fierce reliance on Spotify, and voiced his fears of them ‘leaving’ or disappearing.

“I need this to be universal, I need you to cover everything, and now that you’ve shown me how good it is when you are the only one, you can’t go back.” Boz.

Arianna also said she would feel distraught if separated from her Spotify account and playlists. Others relied on their platform for new music and for new discoveries, while others used it as a working tool or as an archive of music.
“I definitely have periods where I don’t want to listen to anything I’ve heard before and then there are other times where it’s like, ‘Please just give me the 20 songs that I’ve been listening to since I was sixteen’.” Boz.

Depending on his mood Boz would “ask” the platform to either deliver new content, or otherwise give only the nostalgic music of his youth as a comfort to him. Reliance on automated recommendation and curation varied across situations and moods for each user. For instance, Ryo would listen to house or ambient music ‘passively’ at work, using the recommendation or Radio features to create background listening and to leave the curation to the recommender. This would represent the bulk of his listening in terms of time spent, but he felt that this betrayed his ‘actual’ music taste. He didn’t like that the data was ‘muddied up’ like this and there was no way of explaining the difference between the two listening modes to the system.

1.4. Misunderstanding

As in human relationships, despite the recommender system’s ‘uncanny accuracy’ in some cases, misunderstandings between user and platform were also frequent. In continuing the example above, Ryo said, “Yeah 100%, I can’t have my Pavement recommendations knowing that that’s what I listen to!” Like many others, Ryo had thought about and cared about what the system ‘thinks’ of them. Leonard described a troubling example of what might happen when other people used their account causing Spotify to ‘think’ they liked certain things or were a certain kind of person: “Like the thing that would make me upset about that, is if someone put on white supremacist music, and somehow Spotify thought I liked white supremacist music.” Many participants were concerned with how the algorithmic system might misunderstand them, especially through the impact of other people using their account. This de-personalised their account and even had implications for how others (both system and other users) might view them, as the system’s understanding and data profile was considered a reflection of both their taste, and personality. In contrast, when the recommendations were not working for them, some listeners reverted back to discussions of the system as a ‘robot’ which did not have an ability to understand emotions or subjectivity.

Layla described what was a common theme of “their algorithm” being “ruined” and no longer representing them, as someone had used her account at work, listening to music that didn’t represent her taste, and said she felt frustrated at not being able to “wipe away that history”. Similarly, Nina felt that the system had the wrong idea about her personality, her taste and her identity. She said that based off the music that was recommended to her, she believed that the system assumed she was “a Byron mum, boho chic style, and into acoustic guitar”. In these examples, misunderstanding was not just considered to be about their music taste, but also deeply personal. These kinds of misunderstandings about profiles and data are common across many algorithmic settings, but particularly intimate in the music setting.

In other ways many users described a musical life, a history and an identity outside of the app, that couldn’t be accounted for or understood by the system. This included going to see live music regularly, having a classical and/or musical education, or the types of music listened to on other platforms (SoundCloud, Mixcloud, YouTube etc.), as well as songs or artists listened to in analogue formats (LP records). This often contributed to a sense that the system couldn’t really understand them in a wholistic way, and some said that this would be an improvement to the system. Through this sense of alterity, personal attachments can form between user and platform.

These attachments are built through shared and known history, as well as mutual understanding and daily conversations (albeit imagined) between them and the “quasi-other” recommender system. In this way we draw parallels with human relationships, as participants form bonds with the recommender system, or ‘their algorithm’; relationships characterised by trust, reliance and protectiveness (not wanting others to use their account or disturb their recommendations); or betrayal and misunderstanding. Through the above examples, it is clear that personal attachments can form between user and recommender system (and the platform more broadly). However, the characterisations of this relationship (otherness) are shifting and unstable, not only between participants, but also mutable within the participants daily use of the system. In thinking
Don't mess with my algorithm: Exploring the relationship between listeners and automated curation and recommendation on music streaming services

about what the system thinks of them and how that was reflected back in their recommendations, participants pointed to a number of theories for how the system works.

**Theme 2: Theories**

The emergence of theories as a theme demonstrated that many participants were constructing personal theories to understand how the system worked, and to make sense of the content recommended to them. This draws on Bucher’s (2016) work on the *algorithmic imaginary* on social media, and the way that user concepts and theories contribute to their use of the platform. Similarly, this can be related to work on folk theories as leading to idiosyncratic notions of how the platform works (Eslami, et al., 2016). It is important to note that in terms of usability, an understanding of the algorithmic and machine learning processes that underpin the operation of music streaming services is not necessary for their use. Many participants admitted, they only really considered how the platform worked when there was a perceived failure (*i.e.*, bad recommendation). This section explores these theories about how the system works and draws particularly on the observation of participants as they reviewed their personalised and recommended content, specifically their *Discover Weekly* or *For Me* playlists.

2.1. Theories of how the algorithmic system works

Listeners through their extended interactions with the streaming service and algorithmic features would come up with idiosyncratic theories and explanations of how the system works. Marc said:

> “I find the sort of more esoteric I get in terms of artist or genre, decade, style, those sorts of things, the less it tends to rely on what I’ve already listened to in terms of giving me choice And then it’s almost like placing me within that network and then just kind of sending me on spin cycle amongst other people’s like algorithms.” Marc.

Like Marc, others offered different theories of what data or behaviour the system is taking into account in order to make recommendations. Some had technical explanations involving collaborative filtering, and some pointed to the analysis of musical features such as instrumentation. Others posited that the system is taking into account time of day, number of song listens, to whether people listened to or followed popular playlists. This also can be interpreted as what the users found to be the more salient aspects of how or what they were listening to. User’s explanations in these examples show some awareness of an algorithmic logic being used, as well as the output of data that they may contribute towards the analysis and recommendation of songs. As Bucher (2016) notes, participants were also aware of the presence of other users and how their behaviour was also analysed and affected their recommendations. Arianna posited that her recommended content would also be tailored to fit a broader, general community of listeners. Indeed, most participants admitted that they didn’t think about or care for how the system works, until they received a ‘bad’ or inaccurate recommendation.

> “The only time I noticed it’s bad is when it’s really trying to push some sort of something that’s a little more popular. I’ll be listening to a song and then all of a sudden it’s just like, ‘Oh, that’s a complete switch of mood’.” Curtis.

Here Curtis perceived a sort of agenda of the system, that it was pushing certain content out for a reason other than to give accurate recommendations. Other participants would give the system features or affordances it might not actually have, by imagining affordances that go beyond its current algorithmic and computational capacity. For instance, Marc had guessed that there was care taken in the ordering and curation of his *Discover Weekly*, and Leonard pointed to the careful selection of a band’s second or third hit, rather than their number one on their favourite disco *Radio* station.
2.2. Reactions to algorithmically curated content and recommendations

During interviews, listeners were invited to share their Discover Weekly (Spotify) and For You (Apple) personalised playlists, commenting on the recommendations and offering explanations or rationales for why certain content was there. This, again, often was conceptualised as a conversation between listener and ‘algorithm’, system or a persona.

“The times I have checked it, I’ve generally been like, ‘What the fuck? Is this for me? Sure this is for me? Got the wrong person maybe?’” Jay.

As in this example, participants expressed frustration and confusion as to why certain songs were “given to them”, with Marc and Layla also describing asking the system directly why they were recommending that content, or even if there had been “some kind of mistake”.

Observing participants’ reactions to recommended content on personalised playlists provided an opportunity for participants to narrate the way that their taste was reflected back to them, and to comment on how successful this was. Such reactions were telling of their level of satisfaction and trust in the recommender systems, as well as demonstrating perceived differences in how their taste was constructed by the systems, and how they understand music to ‘fit’ together. This notion of ‘fit’ was a common concept brought up by participants, whereby there was a disconnect between the content they were seeking and the content served to them in recommendations or curated content. While responding to their personalised content, as in Jay’s comment above, participants often mentioned feeling that the content they were served didn’t ‘fit’ them, or ‘fit’ the playlist they were listening to, which they expressed frustration or disappointment about.

2.3. Alignment

Misalignment of expectations and recommended context was a recurring issue for participants. During the observations participants appraised the content, often finding songs or artist that didn’t align with their own ideas of what should be on the playlist, and sometimes providing insights into why. Ryo explained:

“I suspect that it is recommending that because that song is popular because it’s on the Fallout soundtracks to the video game, that like re-popularise a bunch of super old music. So, it’s not like insane. But none of those songs would fit on this playlist.” Ryo.

Here Ryo’s theory of why a song was recommended to be added to one of his own playlists is based off his personal conceptualisation and sense-making around how these songs fit together; in this case because they were on the soundtrack to popular video game, Fallout. For Ryo objectively the songs ‘fit’ together, but not on this personal playlist as he has a very clear, and subjective idea of what songs ‘fit’. Ryo described songs as being “informationally adjacent”, as in the artists might link on their Wikipedia pages, but for him they didn’t fit together in terms of the sound, or ‘vibe’, or mood he was trying to curate.

In this way, participant’s understandings could be ambivalent and contradictory. They described recommendations being both too literal and basic, whilst also at times being uncannily accurate and creepy. This depended on their mood, context and intention for their listening session. Often, they could provide explanations for objectively why the system had put the songs into a particular playlist, but many complained that this didn’t ‘fit’.

Other inaccuracies were found in response to Discover Weekly content. Listeners often observed that the platform couldn’t understand or predict their emotions. Marc explained:
“I guess that’s the thing. It can’t really predict how you’re feeling. I don’t think it’s accurate in that sense. So, it might go off everything that you listened to in a week or the last six months or the last 12 months, but, sometimes it seems it can’t really figure out where your taste is going. Therefore, maybe it doesn’t really understand you ... Well, it’s like, here’s a series of things that like a part of the same genre.” Marc.

Carl described listening sessions where he was trying to create an ‘audible journey’, but the recommended content didn’t sync with that. This situation was echoed by other participants, saying that the music was objectively similar but didn’t ‘do’ what they wanted it to, or connect in the way they imagined. Jay described trying out a personalised gym playlist; “I have tried actually, and it was really lame. My idea of working out is really different to Spotify’s”. Carl continued this thread by saying:

“I’ll put it up on a Radio, but I usually find that I have done the like skim and scan beforehand to know whether or not it’s gonna work or whether or not it does hit what I want it to hit.”

Carl.

In some cases this disconnect was actually a positive feature for users, as it put things together in interesting combinations that they themselves wouldn’t have done. Overall, a number of participants were sceptical or distrusting of the system’s capacity to curate a subjective, or emotional listening experience like a human could. “A robot can’t do that” said Layla. Freddie showed a distrust for the technology by saying:

“The algorithms are really impressive, but I don’t want them to be able to do what a human can do ... I’m talking about the progression, having that play after that. And it would just feel right. I don’t think an app should be able to do that as well as the human.”

Freddie.

Though some, like Freddie in this exchange, felt clear distrust and disbelief towards the abilities of the recommender system, others described a personal attachment and a bond that forms between user and system; albeit a bond that can be broken or tested through misunderstanding, misrepresentation and inaccurate recommendations. Listeners also shared their extensive theories for how the service is working, explaining why certain songs were recommended to them, and crucially how and why the system ‘got it wrong’. These theories belie the complex socio-technical relationship forming between user and recommender system, as they attempt to make sense of and understand one another.

**Theme 3: Tactics**

In the final theme, listeners spoke about ways that they attempted to circumvent, navigate or even resist the algorithmic curation and personalisation features on their music streaming platforms. As such, this theme gestures to Certeau’s (1984) notion of tactics, resistance and control, as we see the way that users incorporate this technology into their daily life, and yet find tactics to make it their own, and resist the algorithmic decision making on the platform, as well as resisting the content itself. Not only did most participants have an awareness of the algorithmic system and how it might work, but some employed tactics to train or mould the system to receive the kinds of output they wished for. This training would often involve the use of the Like features, or by adding songs deliberately to certain playlists, or re-listening to particular songs to emphasise their preference for them. Ryo describes actively training the system:

“So if I press ‘like’ on something, I’m think, “this is good, you’re learning and I want you to take note of this”. And so if I’m at work it recommends some music that I like, and I’m distracted, I’ll go back and look through the thing from half an
Don't mess with my algorithm: Exploring the relationship between listeners and automated curation and recommendation on music streaming services

hour ago and go and like those songs because I want it to do the job.” Ryo.

Personal attachments to songs, albums, and even playlists, seem also to be built on effort, work and a history of deliberately training the system to be more personalised (i.e., to become “my algorithm”). Again, participants were very aware of the labour they had put into the system and how the data traces of this labour contributed to their overall taste profile, and thus the identity reflected back to them. While some developed tactics to train the system on what they liked, trying to confirm what the platform had already done, others considered how to avoid being ‘pigeonholed’ by deliberately listening to almost antithetical music to the content they had been recommended.

Marc consciously searched for things that would lead the recommendations in a certain direction, explaining that “it’s like something that you've got to maintain in order to kind of unearth it and make it more frequent through your recommendations”. Marc described this as a kind of “data struggle”, with 80 percent of listening being characterised as daily, regular listening, where recommended content was homogenous and unexciting. The other 20 percent of recommendations he described as being interesting or discovery based, and that he needed to tend to this 20 percent by training the system, or to use Boz’s words, to ‘nudge’ the system in a certain direction, particularly towards new, undiscovered content.

Participants were also acutely aware of the consequences of someone else using their account, and many described having to put in time after this event to adjust or retrain the system back to their preferences. Despite these consequences, there was a sense of social etiquette around letting others use their account:

“If I pass someone the phone in the car, so I would ideally like to press private browsing before I pass it, but that’s too [passive aggressive] for me to do. So, I just try and absorb the hit. And then I’d go home and I listened to The Wipers. I listened to The Raincoats, and listen to like cool music just to try counteract.” Ryo.

Ryo’s desire to use the Private Listening feature on Spotify was echoed by others who used this as a tactic to listen to what they described as ‘guilty pleasures’ or embarrassing music. This demonstrated that, at least on Spotify, some users were very aware of differences around private and public listening, and being observed, or even surveilled by both the platform and other listeners.

3.1. Using recommendation & curation features as tools for discovery

Almost every interviewee mentioned some version of following a chain of recommendations, i.e., going down a ‘YouTube hole’, click hole, worm hole, rabbit hole, which complements the work of McKelvey and Hunt (2019) on flows on content discovery platforms. Many participants described following the recommendations, usually on YouTube but also on Spotify through the Recommended Artists feature.

Jay described using the Radio feature to scout for songs to augment the vibe on a playlist he was making, by using other recommendation features to make his playlist better. Boz described using the recommended artist feature as a kind of rabbit hole, clicking through to discover arcane and strange music:

“I used to do a thing where I would go to one artist, go to their related artist, go to the bottom of the list to find that person who’s like, “I guess we’ll throw this person on”. That’s how I imagine it works, with most related to them. Yeah, go to the bottom list and then doing that like 20 times until you end up with someone and then go from there. You always end up at someone like kind of always really bad, but always at least interesting, like I guess this is what Catalan blue grass music
Some listeners said they wouldn’t like to admit that they found things through algorithmic recommendations. When asked why, Boz said:

“Because I guess, it’s lazy, you got handed it, you didn’t work for it. You didn’t know the cool person. You didn’t go to the gig. You just got given it by the machine. It came out of the slot. You put 11 bucks in a month and the vending machine spits out of a hot slice and then that’s it.” Boz.

This exchange, and others like it, demonstrated that for many listeners a human recommendation or personal discovery had more meaning than an automated recommendation. This hints at the value given to recommended content on music streaming services. Jay said that Spotify had “given” them their two favourite artists of the last five years, which was “pretty significant”, but at the same time he commented that this seems small in comparison to the volume of recommendations made to him by the system. Following on from this perceived value for human, subjective recommendations, Arianna, Jay, Carl and Layla all said that a way to improve the system would be to present the recommendations in a more conversational way. For instance, a recommendation would come with a message like, “hey you might like this” and this would be accompanied by the system offering a form of explanation for why that song was suggested. In this way they wished that the recommendations made by the system could be more ‘organic’, ‘social’ and closer to a human recommendation in the way they are presented. These improvements would make the system align more with the imagined way that they interact with it, that is through conversation and by attributing meaning to the recommended content.

Building off of their own idiosyncratic theories and algorithmic imaginings (Bucher, 2016), listeners attempted to use tactics to train, cajole, mould, or steer the system to either correct perceived inaccuracies or interruptions to their account by others, or to validate correct recommendations, by giving positive reinforcement to the system. Others used recommendation features as a form of radical discovery, going ‘down the rabbit hole’ in search of weird, and arcane songs and artists. This relates back to Ihde’s (1990) notion of alterity, as the users were aware of the quasi-otherness of the platform and attempting to exploit this through tactics to gain better, more accurate recommendations. In other cases, the system was described as simply a tool for previewing, listening or for discovery. This raises compelling questions of agency and control on the platform, as the system can function as both a tool for users, as well as an agentive force. Overall, this theme demonstrated the desire of listeners to actively influence the nature of the recommendations they were offered by the streaming service, and an attempt to correct misunderstandings about them and their taste.

Conclusion

In line with McKelvey and Hunt’s (2019), observation that “although agency is distributed on content discovery platforms, it is not distributed equally”, a dynamic tension emerged for users between agency and reliance on music streaming platforms and their recommender systems. The recommender system could function as a both a tool and an agent in different scenarios and contexts. The system was agentive as it is making decisions based on listener data and music analytics to algorithmically curate and recommend music for listeners. As such, music recommender systems and platforms are the nexus of human and machine knowledge of music. Human actions such as listening, skipping, sharing, provide constant machine learning training data. The recommendations and automated curation thus shape and frame listening sessions, raising interesting questions about who is the expert on music taste — listeners themselves or the algorithms that recommend music to them? This relationship is important, as participants ‘allow’ the system
into their most intimate spaces; music to accompany sleep, work, gym sessions, breakups, dinner parties and so on. As in other algorithmic settings, it is crucial to understand ‘who’ we allow to make proxy decisions for us, judging carefully how well they make those decisions, and crucially also asking how those decisions are made at a technical level. Listeners appear in some way cognisant of this relationship and this algorithmic framing, and at times will relish these affordances, and at other times push back. The relationship between the user and the system is dynamic, and shifts daily depending on mood, music and discovery needs, context and the listening mode.

Whilst this study focused on the music recommendation context, a number of findings can be related to other algorithmic settings. Listeners anthropomorphised the recommender system, imagining a conversation between them and “the algorithm”. As in Bucher’s (2016) work on folk theories, participants had idiosyncratic theories and tactics for how to understand and navigate the system. Like in other algorithmic contexts, on music streaming platforms behavioural data is taken as proxy for human emotions and intentions, arguably creating a feedback loop (Morris and Powers, 2015). Finally, despite the utility and ‘uncanny accuracy’ of music recommendation and curation on music streaming services, as in other algorithmic settings, listeners are not afforded much control or transparency over algorithmic filtering, decision-making and automated curation. Holding a position of power and privilege, music streaming services mediate, via commercial and algorithmic logics, the human experiences of music discovery, curation and recommendation.

In this study participants tended to be very interested in music and technology, as they were recruited as daily users of the platform. Some participants worked with music or were classically trained musicians, which would have an influence on how participants use and engage with music streaming services, and with music itself. Future research should look to other demographics, with a broader scope for participants interest in music, and technology. Building on these notions of the relationship that forms between listener and system, future work could move toward observation of task-based or contextual listening sessions, especially to expand the research into tactics employed in response to music recommendation. Opportunities exist to further explore the different listening and discovery modes, as well as the reliance on recommender features across other streaming platforms [5].

About the authors

Sophie Freeman is a Ph.D. candidate in the School of Computing & Information Systems at the University of Melbourne, Australia.
Direct comments to: sophie [dot] freeman [at] unimelb [dot] edu [dot] au

Martin Gibbs is a professor in the School of Computing & Information Systems at the University of Melbourne, Australia.
E-mail: martin [dot] gibbs [at] unimelb [dot] edu [dot] au

Bjørn Nansen is a senior lecturer in media and communications at the University of Melbourne, Australia.
E-mail: nansenb [at] unimelb [dot] edu [dot] au

Notes

1. In this case, our research is speaking to the socio-technical relationship that forms between user and system. However, this paper also aligns with the conceptualisation by Goldschmitt and Seaver (2019) of algorithmic systems more broadly as complex socio-technical apparatus, as they are hybrid of human and machine work.


5. Spotify was the dominant platform among participants \((n=13)\), followed by Apple music which was used exclusively by two participants. However, many participants used other services in different situations, such as Bandcamp, SoundCloud, YouTube and YouTube Music, as well as music recognition app, Shazam. All of these services are worthy sites of future analysis.

References


Don't mess with my algorithm: Exploring the relationship between listeners and automated curation and recommendation on music streaming services

volume 12, number 3, pp. 233–267.


Don’t mess with my algorithm: Exploring the relationship between listeners and automated curation and recommendation on music streaming services

7, number 2, pp. 95–116.


K.L. Turner (editor), 2016. This is the sound of irony: Music, politics and popular culture. London: Routledge.


Editorial history

Received 7 September 2021; revised 11 November 2021; accepted 30 December 2021.

Copyright © 2022, Sophie Freeman, Martin Gibbs, and Bjørn Nansen. All Rights Reserved.

‘Don’t mess with my algorithm’: Exploring the relationship between listeners and automated curation and recommendation on music streaming services
by Sophie Freeman, Martin Gibbs, and Bjørn Nansen.
Don't mess with my algorithm: Exploring the relationship between listeners and automated curation and recommendation on music streaming services

First Monday, volume 27, number 1 (January 2022).
doi: https://dx.doi.org/10.5210/fm.v27i1.11783