The Dynamics of Exploration on Spotify

Lillio Mok[1], **Samuel F. Way**[2], **Lucas Maystre**[2], **Ashton Anderson**[1]

[1]Department of Computer Science, University of Toronto
[2]Spotify Research
{lillio, ashton}@cs.toronto.edu, {samfway, lucasm}@spotify.com

Abstract

Digital media platforms give users access to enormous amounts of content that they must explore to avoid boredom and satisfy their needs for heterogeneity. Existing strands of work across psychology, marketing, computer science, and music underscore the importance of the lifecycle to understanding exploratory behavior, but they are also often inconsistent with each other. In this study, we examine how users explore online content on Spotify over time, whether by discovering entirely novel music or by refreshing their listening habits from one time frame to the next. We find clear differences between users at different points of their *off-platform* lifecycles, with younger listeners consistently exploring unknown content less and exploiting known content more. Across their *on-platform* histories, users also explore in bursts by following seasonal cycles and exploratory phases. We also find that these patterns of exploration do not translate to other notions of heterogeneity like diversity; notably, younger listeners are more diverse in their consumption despite exploring less. Exploration and diversity thus capture different ways in which people find variety, potentially accounting for the inconsistencies in existing work. Together, these nuanced dynamics of exploration suggest that online platforms may be better poised to support users by incorporating different measures of heterogeneous consumption.

Introduction

People crave heterogeneous experiences and dread boredom (Kidd and Hayden 2015; Zuckerman and Neeb 1979). To assist users in satisfying these needs, digital streaming platforms provide them with access to massive amounts of content along with recommender systems to supplement their interactions with the platform. Ideally, these systems should help users find variety when they are most receptive (Schedl and Hauger 2015; Zhang et al. 2012) and when they become tired of repetition (Benson, Kumar, and Tomkins 2016), thus motivating the need to understand heterogeneity-seeking behaviors online. For example, existing work illustrates how people explore individual pieces of content through discovery (Datta, Knox, and Bronnenberg 2018), novelty (Zhang et al. 2012; Kapoor et al. 2015), and repeat and non-repeat consumption (Ratner, Kahn, and Kahneman 1999). These behaviors can be highly variable, depending on seasonality (Park et al. 2019), the environment (Way et al. 2019), personality (Anderson et al. 2020b), and user demographics (LeBlanc et al. 1996).

However, people also often explore in dynamic ways that change over extended periods of time. The importance of users' *lifecycles* in understanding variety-seeking behaviors is therefore a recurring theme in literature spanning psychology (Novak and Mather 2007; Nicklaus et al. 2005), marketing (Holbrook 1993; Datta, Knox, and Bronnenberg 2018), and computer science (Danescu-Niculescu-Mizil et al. 2013; McAuley and Leskovec 2013). *Online lifecycles* broadly follow a trajectory in which discoveries of new content are initially common and become increasingly rare, leading to increased boredom (Benson, Kumar, and Tomkins 2016). People also display varying preferences for heterogeneity throughout the offline, *human lifecycle*. Old age leads to more nostalgia (Batcho 1995), less openness (Costa Jr et al. 1986), and the derivation of less meaning from novelty (Carstensen, Fung, and Charles 2003). Nonetheless, others also find that reduced social pressure to conform could encourage exploration later in life (Bourdieu 1984).

These inconsistent observations are further reflected specifically in research on music, in which findings indicate conflicting relationships between lifecycles and musical exploration. While openness at earlier ages may suggest increased exploration in younger listeners (Ferwerda, Tkalcic, and Schedl 2017; Anderson et al. 2020b), expanding tastes and cultural preferences over time could lead conversely to exploration in older listeners (Park et al. 2015; Peterson 2005). Thus, despite the emphasis placed on the lifecycle when characterizing exploration, empirical gaps remain in our knowledge of how users actually move between pieces of online musical content across their lifecycles.

Against this backdrop of existing research, the present work contributes a large-scale empirical study of how users explore online musical content throughout their lifecycles. We guide our work through the following research question:

- **RQ:** *How do users explore individual pieces of online content over their on- and off-platform lifecycles?*

To address this question, we conduct a longitudinal study of how 100,000 US-based listeners on Spotify explore musical content. Our work is organized into three main sections. First, we study general patterns in exploration as a function of their physical lifecycle and age over two time scales. In the long term, we measure how much and how often users add new content to their listening repertoires, and, in the short term, we measure how content cycles into these repertoires. Second, we investigate whether exploration is evenly distributed over their on-platform histories, or if it occurs in phases. Finally, we contrast notions of exploration and taste diversity by comparing the rates at which users explore and the diversity of the content that they encounter.

Overview of results. We find substantial between-user variation in exploration across users' off-platform lifecycles. In agreement with one strain of existing musicological work, older users are more likely to explore—they discover more pieces of previously-unknown content across the trace, and also have more content turnover on a weekly basis. Younger users systematically explore less, especially older content. This provides large-scale evidence supporting the notion that variety-seeking behaviors increase as people age.

Furthermore, we observe temporal variance in exploration within users' on-platform lifecycles. Listeners explore in phases, such that they can be in either relatively exploratory or exploitative periods, and also follow an annual cycle of discoveries around the December holidays. This suggests that exploration is closely associated with transient and seasonal musical preferences during the on-platform lifecycle.

Finally, we find that our content-agnostic exploration metrics are poorly correlated with content-aware metrics like consumption diversity. Younger listeners are relatively more diverse despite being less exploratory, whereas older users are less diverse despite being more exploratory. Exploration and diversity thus capture distinct notions of heterogeneous consumption, which accounts for the conflicting theories of preference ossification over listener lifecycles.

Background

The present work builds upon three bodies of literature, which we survey in this section. We first describe theories and mechanisms underpinning the differences in variety-seeking behavior across age groups. We then consider evidence pointing to age-dependent needs for musical variety. Finally, we delineate how user preferences for heterogeneity are studied in the context of online platforms.

Theories of heterogeneity and lifecycles. There is good reason to believe that people change their preferences for variety as they age, due e.g. to changes in personality traits. For example, some observe older subjects' openness to experience to decline as preferences "set-in" (Roberts, Walton, and Viechtbauer 2006; Srivastava et al. 2003). Others find openness to remain flat across most of one's middle years, with significant dips and peaks in early and late life (Soto et al. 2011). Furthermore, younger people are prone to being bored (Vodanovich and Kass 1990), and need to be curious for cognitive development (Gibson 1988; Kidd and Hayden 2015). Socioemotional Selectivity Theory argues that, as they age, people become more focused on deriving meaning from learned experiences and less on novelty (Carstensen, Fung, and Charles 2003). This literature suggests that preferences tend to ossify over the human lifecycle.

Variety in music. These theories between age and preference for variety also translate to music. Cognitively, music serves as a means of alleviating boredom (Hargreaves 1982) and satisfying epistemic curiosity prevalent in younger listeners (Zuckerman and Neeb 1979). Some also postulate the concept of "open-earedness", suggesting that young listeners might be more aesthetically diverse (Hargreaves 1982). Affectively, music enables mood regulation in adolescents such as helping to manage stress (Saarikallio and Erkkilä 2007; Saarikallio 2011) and may have emotionally therapeutic properties (Tervo 2001). In contrast, older listeners may have more stable, if not enhanced, existing capabilities for emotional regulation (Carstensen, Fung, and Charles 2003). Socially, music allows youth to build an identity amongst their peers (Tarrant, North, and Hargreaves 2002).

Empirical trends in musical tastes appear to both agree with and dispute these mechanisms. On the one hand, the demonstrable fluctuations in personality traits as listeners age are likely to mirror similar fluctuations in musical tastes (Anderson et al. 2020b; Ferwerda, Tkalcic, and Schedl 2017). Openness has been shown to correlate with higher rates of music discovery and higher diversity of tastes (Anderson et al. 2020b). The decline of openness through the aging process would suggest a corresponding decline in listening variety. However, other studies point to a rise in the number of liked genres from adolescence into adulthood, reaching a peak between the ages of 45-55 (Hargreaves and Bonneville-Roussy 2018). Similarly, there is evidence that preferences follow a U-shaped distribution over time, narrowing during adolescence and rising into adulthood (LeBlanc et al. 1996).

Variety-seeking behaviors online. Beyond music, our work builds upon a long line of research investigating online heterogeneous consumption over the user lifecycle. Prior work has found that tastes evolve over time as users gain expertise (McAuley and Leskovec 2013), that users can become bored as their consumption becomes stale (Kapoor et al. 2015; Benson, Kumar, and Tomkins 2016), and that linguistic behaviors change most rapidly near the beginning of the user lifecycle and stabilize later on (Danescu-Niculescu-Mizil et al. 2013). A key notion in this literature is *diversity*—the coherence of the items consumed by individual users. The generalist-specialist (GS) diversity score that we use was originally applied to Reddit data, where specialists and generalists can predictably engage more with specific content or the broader platform (Waller and Anderson 2019). In the recommendation systems literature, there have been several attempts to incorporate notions of novelty, diversity, or serendipity (Schedl and Hauger 2015; Zhang et al. 2012; Schafer, Konstan, and Riedl 2002; Zhang and Hurley 2008).

On music streaming platforms specifically, listeners follow musical paths according to changes in taste over time (Moore et al. 2013). Recent work suggests that the adoption of streaming services promotes novelty-seeking behavior, with older users discovering more novel content than younger users (Datta, Knox, and Bronnenberg 2018). On Spotify, listener satisfaction as indicated by conversions to paid accounts and retention on the platform was shown to correlate strongly with diverse listening (Anderson et al. 2020a). Characterizing heterogeneous consumption over time may therefore have substantial impact on streaming platforms; experimental work on podcasts specifically find tensions between diverse listening and and user engagement (Holtz et al. 2020).

Relation to this work. This rich body of literature underlines the motivation for our research question. Whether peoples' tastes ossify as they traverse their offline lifcycles remains empirically unanswered across many listeners. While some evidence points towards narrowing preferences in older consumers, others suggest that the acquired familiarity with different content types leads to broader tastes. Furthermore, in the context of on online platforms, granular, content-agnostic metrics like exploration and coarse-grained, content-aware metrics like diversity have been introduced separately without comparison. Our work identifies the disconnect between the two through the lens of online lifecycles, and shows that exploration and diversity measure different aspects of how variety-seeking behaviors evolve.

Data

We study Spotify, an online platform for streaming music with over 50 million available songs. Its services are available either for free or on a monthly paid subscription ("premium"). On the free version, users are served advertisements interspersed into their listening and have some restrictions over on-demand features. Premium users do not have these restrictions. Both versions are available as native applications on mobile, desktop, or Internet-of-Thing devices, and also function through modern Web browsers.

Our main data set consists of the listening traces of over 8B unique listening events from 100K US-based users, spanning 4 years between 2016 and 2019. We restrict our analysis to premium users who remain active and subscribed in each month, and consider only the users whose self-reported ages are between 18 and 65 years at the start of the trace. Streams that lasted under 30 seconds are discarded.

Although seemingly non-representative, our inclusion criteria of 4 active years is necessary in order for us to consider the *complete*, longitudinal dynamics of exploration over the lifecycles of *many* users. In comparison, a popular Last.fm dataset contains only 275 users with self-reported ages[1]. Furthermore, we found our dataset to be comparable to a representative sample of premium Spotify users without activity constraints in the same period. This includes age composition (21% 18-24, 51% 25-34, 18% 35-44, 7% 45-54, 3% 55-64 in our dataset vs 13%, 52%, 21%, 9%, and 4%) and lifetime track-to-stream ratios per group (16%, 22%, 26%, 25%, 26% vs 17%, 23%, 27%, 28%, 30%). Activity levels (30 vs 27 mean streams per day) and genres consumed (24 vs 24 on average) were also similar. It is therefore very likely that our results will generalize to broader populations.

Programmed content on Spotify is defined as content driven by Spotify's platform, e.g. personalization and recommendation, algorithmically generated radio stations, or curated playlists. This content encapsulates whenever a user initiates a listening session in which subsequent songs are determined without user interaction. *Organic* content includes tracks whose streaming is determined by the user or other users, such as playlists or songs for which the user directly searched. Note that these definitions are applicable broadly to any online platform whose content is driven this way, be it through hand curation or recommendations.

[1] Available at http://ocelma.net/MusicRecommendationDataset/lastfm-1K.html

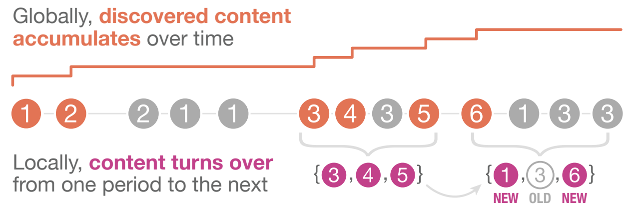

Figure 1: Schema of exploration at different time scales.

We also use an embedding of tracks on Spotify later in this study. This embedding considers user-curated playlists as "documents" in which each track is a "word", and uses the word2vec continuous bag-of-words model (Mikolov et al. 2013) to place these tracks in a 40-dimensional musical space. Tracks that are often co-located in playlists will be closer together in the space. This method has been demonstrated in previous work to yield a metric space that is congruent with how users listen to music on Spotify (Anderson et al. 2020a). For consistency, we use track vectors from a single embedding trained at the start of 2020.

Defining exploration. We primarily use two metrics of exploration: one measuring exploration of globally-novel content, and one measuring exploration of locally-novel content (see Figure 1). Considering a user's online lifecycle *globally*, exploration materializes in discoveries—when they encounter new pieces of content for the first time. Formally, the ith track streamed by the user s_i is a discovery if $s_i \notin \{s_j : j < i\}$. Similar methods have been used in work on repeat consumption and boredom (Benson, Kumar, and Tomkins 2016). Note that if we inspect a track's cth stream instead of its first, we move from measuring exploration to exploitation in the form of repeat consumption for high values of c. We discuss exploration versus exploitation below.

On a *local* scale, exploration is reflected by content cycling in and out of a user's repertoire between short-term time frames. This includes new discoveries and already-known content from earlier encounters, although both serve as turnover that freshens the user's immediate online experiences. Formally, the metric we use is each user's weekly, stream-weighted track turnover rate. For a user's stream sequence S_i and unique tracks T_i in the ith time window (i.e. one week in this analysis), we identify tracks $T_i^+ = \{t : t \in T_i \setminus T_{i-1}\}$ for $i > 1$. These are the tracks that cycled or were inserted into the user's consumption during the ith window. The corresponding number of incoming streams in i is $|(s : s \in S_i \wedge s \in T_i^+)|$. After normalizing by $|S_i|$, this yields the fraction of incoming streams that cycled into

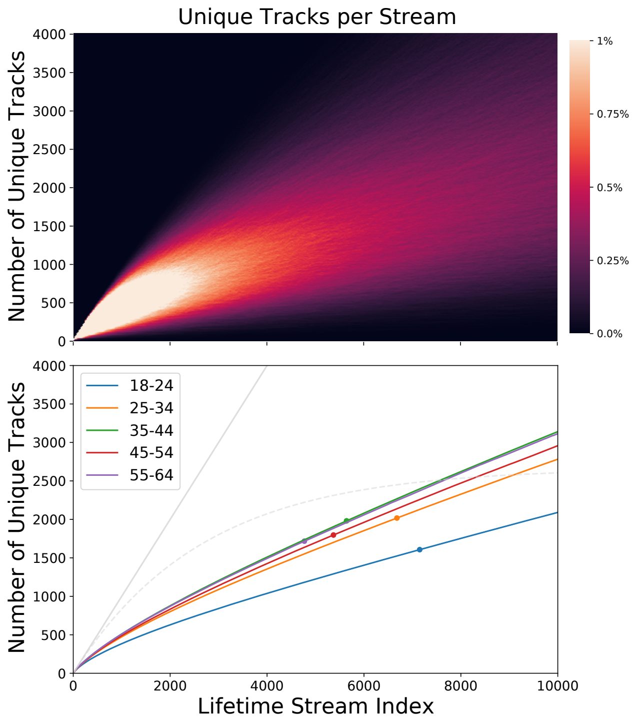

Figure 2: *Top*: distribution of users over the number of unique tracks listened, conditioned on stream count. *Bottom*: expected number of unique tracks per stream, per age group. Mean stream index after one year shown as dots. Identity (i.e., each stream is a discovery) shown in solid grey; draws-with-replacement curve shown in dashed grey.

the window[2]. Variants of this metric have been used to, for instance, quantify novelty preferences (Kapoor et al. 2015).

Exploration and User Lifecycles

What is the relationship between exploration and users' off-platform lifecycles? In this section, we examine how three aspects of how users at different ages and points of the lifecycles explore online content. First, we analyze cumulative global discoveries across traces generated by users of different ages. Second, we measure how different types of discoveries vary depending on user age and content age. Finally, we investigate the generalizability of our results to local exploration in the form of content turnover.

Cumulative Global Discoveries

On a global, long-term scale, how often do users find and consume new pieces of content? To address this question, we operationalize global exploration as *discovery*, when a piece of content is first consumed by a user across their listening history. We measure the number of discoveries against the number of streams. This is equivalent to the cumulative number of unique tracks at each point of a user's trace.

Figure 2 depicts the distribution of users depending on their aggregate unique track count every 10 streams in their trace (*top*). This is then averaged over each user age group (*bottom*). We present two baselines. First, the identity line is shown in solid grey. This would reflect a hypothetical scenario where users constantly explore and discover on each listen. Another baseline models streams as drawing items from a set of 2668 tracks (the mean number of unique tracks after 10k streams) with replacement, shown in dashed grey. This models a scenario where users listen at random with undiscovered tracks eventually being exhausted.

Two patterns are apparent. Firstly, there is a vast amount of variation between users, as depicted by the broad tail of Figure 2 *(top)*. We find this to be attributable in part to the different exploration patterns between different age groups in Figure 2 *(bottom)*. The youngest group has noticeably fewer unique tracks per stream than other groups. This provides preliminary evidence indicating that younger users discover less than older users on a global, long-term scale. They listen to fewer distinct tracks despite having more streams in their trace. In other words, a smaller set of tracks is needed to build their streaming histories than older users. The remaining groups are more likely to make discoveries, which increase in frequency according to their age[3].

Secondly, users' listening behaviors are generally consistent with heavy repeated consumption. With the vertical axis shown at a much smaller scale than the horizontal one, the identity line is very steep on this plot. In contrast, repeat listens outstrip discoveries in users' actual listening traces, and thus yield flatter discovery curves. However, the second baseline indicates that discoveries do not occur as if users were front-loading exploration, as if they were randomly sampling a fixed-sized repertoire of music. Instead of reaching a plateau, discoveries are made more gradually and continue to grow with more streams. A similar perspective is that exploratory activity in music listening appears to be additive. The sublinear discovery curves suggest that repeat consumption of old discoveries persists over time and dilutes new ones. If recently-discovered content subtracts from and replaces older tracks, one would expect steeper curves[4].

Exploration versus exploitation. Content consumption on platforms such as Spotify ranges from exploration, when users encounter content they haven't previously consumed, to exploitation, when users re-consume content they have interacted with many times in the past. We have seen that

[2]We also considered *outgoing* streams, i.e. $T_{i-1} \setminus T_i$, but found this to yield qualitatively similar results. For simplicity we focus on incoming streams, the local analogues of global discovery.

[3]Note that the 35-44 age group disobeys age ordering; see Appendix A for an analysis using the genre of children's music.

[4]We test the persistent repeat consumption of discovered content using a power-law fit. Formally, we fit $y = ax^k + b$ to each of the age groups in Figure 2, where y is the expected cumulative number of discoveries, x is the cumulative number of streams, and a, b, k are parameters. We include x up to 100k streams. This yields exponents k ranging from -0.50 to -0.56 ($r^2 > 0.99$), indicating that cumulative repeat consumption grows much more quickly than cumulative discoveries. However, because this exponent is non-zero, discoveries do not plateau as if users will eventually finish exploring a fixed set of items.

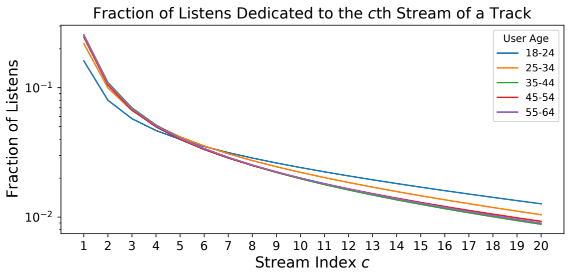

Figure 3: Number of times users listened to tracks for the cth time, normalized by total streams per user.

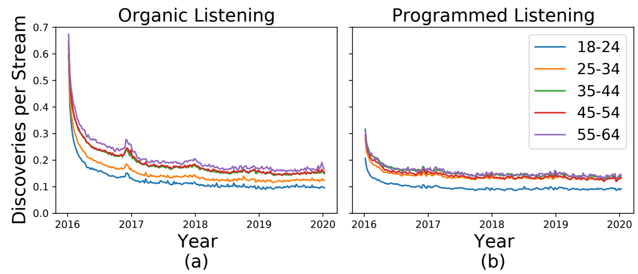

Figure 4: Expected rate of discoveries found from (a) programmed or (b) organic content as a fraction of all listening.

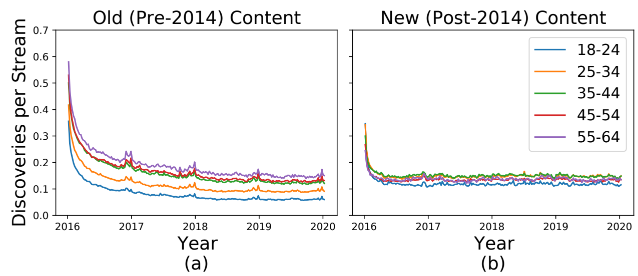

Figure 5: Expected rate of discoveries found from (a) older or (b) newer content as a fraction of all listening.

younger age groups tend to explore less – does this imply that they exploit more in their listening?

For every stream in a user's history, we calculate how many times c they have listened to the song up until the present moment. For discoveries, $c = 1$ since it is their first stream, and $c = 10$, for example, denotes the 10th time they have listened to the song. To answer the question, "How often are users exploring versus exploiting?", we then measure every user's distribution over c. Users who spend more energy exploring novel content will have distributions skewed towards low values of c, while users who repeatedly listen to the same music will be skewed towards high values of c.

In Figure 3, we show these distributions grouped by age bucket. Consistent with our previous results, we observe that younger users explore less and older users explore more, as reflected by the order of the lines for low values of c. We also observe that younger users exploit more than older users as the line ordering is reversed for higher values of c. Thus, on Spotify, people systematically vary in how they approach the explore-exploit trade-off as a function of their age; older users tend to explore and younger users tend to exploit.

This analysis illustrates that the nature of exploration is different for younger and older people. Although younger people explore less, more of the tracks they do explore are eventually converted into favorites that they later reconsume. Additionally, in Appendix B we present an analysis showing that younger users take much less time than older users to convert discoveries into tracks with many repeat listens. These findings together show that exploratory behavior differs across age groups, with younger listeners exploring less often, exploiting more often, and more quickly converting explorations into their top music.

Because this analysis also shows that many discoveries are not repeated and stay at $c = 1$, we continue focusing on $c = 1$ in the present work. Consider a listener who hears their favorite artist's oldest release for the first time, which they realize they dislike. Non-repetition does not detract from their excursion into an unexplored musical locale.

Different Kinds of Discovery Over Time

Our results measuring cumulative discoveries illustrate how users explore relative to their listening history, but do not capture patterns across different types of discovery. Thus, we further compare *programmed* and *organic* content, as defined previously, and *old* and *new* content. The latter is defined in a simple, absolute way: tracks released before 2014 are considered old, and all other tracks are considered new[5].

Additionally, how often discoveries are found may be contingent on when users explore. To capture this information, we consider the temporal dynamics of discovery in this section. We split users' listening trace into weekly buckets, and compute the fraction of streams that were discoveries in each bucket. Given the ith week, the stream sequence S_i during that period, and the corresponding unique tracks T_i streamed in S_i, we compute the number $|\{t : t \in T_i \setminus \bigcup_{j<i} T_j\}|$ and normalize by $|S_i|$. We restrict the raw discovery count to the aforementioned types of content (organic and programmed, old and new) while maintaining the same normalizing constant $|S_i|$. This enables us to interpret the types of discoveries that contribute to overall listening.

Figures 4 and 5 depict this weekly fraction of discoveries across our data trace, split by mode of listening and content age respectively. Again, we find substantial differences in exploratory behavior between age groups. Younger listeners discover fewer pieces of older content and are less likely to do so in an organic manner. Furthermore, these age differences persist systematically across calendar time. We also experimented with raw discoveries (i.e. without normalizing by $|S_i|$) and found the same trends.

Thus, age differences across different time slices and different types of discoveries echo the patterns in cumulative

[5] We experimented with alternative definitions, such as the age of tracks when they are individually discovered by individual users, but found near-identical results.

discoveries (Figure 2). Indeed, notice that 25–34 year-olds, the second youngest group, also discover less than older users across the trace. They unsurprisingly sit between the youngest and older groups. However, the age groups collapse together when measuring discoveries of new content. This shows that younger users have exploratory preferences for recent content but not older content[6].

Secondly, the discovery rate of older content also peaks around holiday seasons, which we find to persist even without activity normalization. Although they are present in younger users, these peaks are especially striking in older users. Thus, music listening habits are sensitive to seasons, with older users exploring songs associated with certain times of the year at the corresponding times. To validate this, we further checked for exploration of holiday-specific genres during the Christmas season, and found age-dependent patterns mirroring discoveries of old content in December in Figure 11 (a) (Appendix A). Thus, these patterns can be understood as another facet of the interaction between user age, content age, and how much is being discovered – particularly of younger users' indifference towards older content.

Finally, relative to user activity on the platform, the discovery rate of content decreases over the first two years of the trace for older content and tracks listened to in an organic manner. We find that the declining prevalence of discoveries is not due to users front-loading their discoveries, such that they stop exploring once a fixed-size corpus is found. We conducted identical analyses to Figures 4 and 5 without normalizing by $|S_i|$, and found raw discovery rates to be stable after the first year – *except* discoveries of old content, which constantly decreases in volume[7].

Instead, users discover older tracks less over time as if content were gradually becoming stale. This is also the main contributing factor to the decreasing organic discovery rates in Figure 4 (a), as 31% of the organic streams in our trace are of older content. In comparison, newer, programmed content is discovered at a flat rate, consistent with users constantly following new track releases and listening to personalized radio. This corroborates our earlier findings: the cumulative discoveries in Figure 2 do not reach a flat asymptote, which would indicate users exhausting their capability for discovering unknown content. Rather, there is a dilution effect: repeat consumption of a growing pool of discoveries outpaces but does not replace exploration of the undiscovered.

Note that age has been a key construct throughout our results. Appendix C presents a robustness check to ensure that it is not confounded by user tenures on the platform.

Local Exploration and Turnover

Thus far, we have found pervasive age differences in the long-term, global discoveries of novel content. To what extent do these differences also generalize to short-term, local exploration? To address this question, we investigate *content turnover* (see schema in Figure 1). On this short-term scale, exploration occurs when pieces of content cycle in and out of users' immediate time windows.

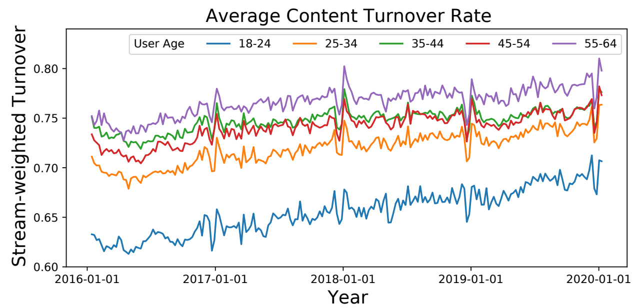

Figure 6: Expected weekly turnover rates over time.

Figure 6 illustrates the average weighted turnover rate per weekly window, per user age group. We see clear patterns in local exploration mirroring those of global exploration. Firstly, different age groups also explore at different rates at the local scale. Younger users have lower rates of weekly content turnover, relative to their activity, across all years. In contrast, older users are more likely to overhaul a significant portion of their listening habits between each week. Thus, we again find evidence indicating that younger users explore less, this time on the local scale.

Another point of interest is the annual cyclic patterns, especially in the older groups. Recall the spikes in discoveries during Christmas (Figure 5 a). Here, local turnover peaks around the holidays with an additional dip during the week of Christmas itself. This suggests that users change their listening habits gradually going into December holidays, during which they hold their set of seasonal music in memory and explore very little. After the festive period, they spike in turnover as they discard these songs and return to non-holiday listening. Thus, although the annual cyclic dynamics differ between local and global exploration, both types of exploration display clear seasonality patterns[8].

Temporal Heterogeneity in Exploration

Our preceding investigation speaks to clear distinctions in exploration *between* different user groups at different points of their *off-platform* lifecycle, operationalized as age. However, this does not capture temporal variance in exploration *within* users' individual *on-platform* lifecycles. Is exploration evenly dispersed across a user's on-platform listening trace, or do listeners explore in bursts?

To address this, we split individual users' histories into weekly time windows, and calculate each user u's discovery rate for each window i. We restrict this analysis to weeks in

[6]See Appendix A for a genre validation. This also reflects the time-to-conversion analysis in Appendix B, in which recent discoveries convert sooner for young listeners.

[7]Another explanation for decreasing discoveries is that our definition implies all tracks at the head of the trace are discoveries. However, this occurs only for the very beginning of traces. The sharp decline in new content discoveries (Figure 5 (b)) indicating this phenomenon disappears after the first half of 2016.

[8]This is again validated using Christmas genres in Appendix A. Note that turnover increases over time across all groups. This corroborates Figure 2. As users' corpora of known music grow with continuous discovery and exploration, more turnover is needed for continuous repeat listening and exploitation of known songs.

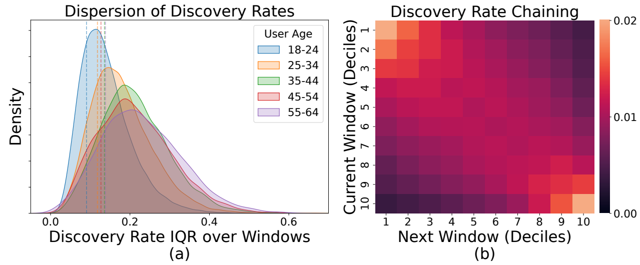

Figure 7: Distribution over within-user inter-quartile range of weekly discovery rates with permuted medians shown dashed (a); probability of weeks with high and low within-user discovery rates conditioned on previous week (b).

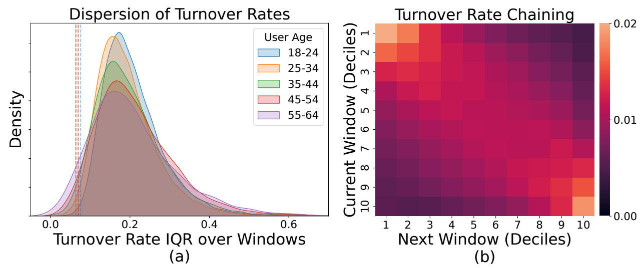

Figure 8: Distribution over within-user IQR of weekly content turnover rates (a); probability of weeks with high and low turnover rates conditioned on previous week (b).

which users were active, i.e. weeks in which their total number of streams exceeded their personal weekly first decile. In order to inspect how individual discoveries are spread across windows, we measure the inter-quartile range (IQR) of weekly discoveries per user. Consider two users for whom 10% of streams are discoveries. A user with an even distribution of discoveries may discover at a 10% rate each week, whereas one with bursts may discover at a 20% rate in alternating weeks. The second user's weekly IQR would be significantly higher than the first (0 vs. 20 percentage-points).

Figure 7 (a) illustrates the distribution of weekly discovery IQRs, separated by age groups. The median discovery IQRs for the youngest to oldest groups are respectively 0.13, 0.17, 0.20, 0.20, and 0.22. To form a baseline, we also measured discovery rate IQRs with permuted traces instead. For each user, we shuffled the track identifiers for each stream in their history, preserving the timing and order of each listen but shifting when discoveries occur. In this case, we obtain median IQRs of 0.09, 0.12, 0.14, 0.13, and 0.14 (dashed lines). The lower weekly discovery IQRs under permutation were significant across every age group ($p < 0.0001$ for each group using Holm-corrected Wilcoxon tests), and only 12% of users have lower IQRs than their permuted selves.

This suggests that users' discoveries are concentrated in windows with high discovery ratios, whereas many of their other windows have few discoveries. Discoveries are thus *skewed* towards a subset of windows, which indicates that users explore in bursts. Because the permuted traces contain identical amounts of discoveries and unique tracks, this further shows that the phases are not an artifact of the variance in aggregate discoveries shown by Figure 2.

However, despite the uneven concentration of discoveries into windows with high and low rates of exploration, the exploration-heavy and exploration-light windows themselves may be evenly disperse. For example, consider a user who explores exclusively for the first half of their trace and repeatedly listens for the second half, and a second user who alternates between exploration and exploitation in adjacent windows. Both would have similar windowed discovery IQRs, even though the latter explores more homogeneously.

To test for this, we first rank each user's windows by their their within-user discovery rates. The weeks are bucketed into deciles according to this ranking. So, even if a user is not particularly exploratory compared to the rest of our sample, spikes in their personal exploration rate will still be identified as being in higher buckets (e.g., weeks in a user's 10th decile). Finally, we sample 10 pairs of consecutive windows per user (e.g. the 61st and 62nd week of a given user's trace).

Figure 7 (b) illustrates the distribution of (u, i) pairs, according to how exploratory user u was during windows i and $i + 1$ in this bucketed manner. Note that, if a user is in a peak of discovery during a given window, they are extremely likely to continue this exploratory streak into the next window. Conversely, if they are in a discovery trough, they are also unlikely to discover in the following window. This reinforces the notion that exploration is heterogeneous over time and points to *chaining* behaviors in exploration.

Whereas the preceding section showed that there is heterogeneity *between* users in terms of exploration, the present finding speaks to heterogeneity *within* the users themselves. The streakiness of discoveries suggests that exploration is not evenly-distributed over time. Instead, discoveries are unevenly distributed into weeks of exploration and exploitation, and discovery-heavy weeks further occur in chains.

Turnover Phases. Although global exploration occurs in phases and chains, is the same phenomenon reflected in local exploration? We similarly apply the same windowed analysis to content turnover to check for temporal heterogeneity in how users explore content at a more local level. Figure 8 illustrates local variants of the global exploratory phases in Figure 7 by using identical windowing functions over turnover rates.

In (a), we find median turnover IQRs of 0.19, 0.18, 0.19, 0.20, and 0.19 for the youngest to the oldest age groups. By permuting each user's trace, we find the null median turnover IQRs to be 0.08, 0.06, 0.06, 0.07, and 0.07 (dashed lines; $p < 0.0001$ for each group using Holm-corrected Wilcoxon tests). Analogously to Figure 7, the IQR of weekly turnover rates reflects how skewed content turnover is between weeks of high and low turnover density. Compared to their permuted selves, actual users group turnover into weeks of concentrated local exploration, while turning over little content in other weeks. Only 3% users have lower IQRs than their permuted selves.

In (b), we find near-identical evidence of chaining effects in local exploration. If a user is undergoing a burst of content turnover in a given week, they are likely to also continue this spike in the next week; the same holds true for weeks with unusually low exploration. Therefore, like global exploration, local exploration also occurs in phases and chains.

Thus, as opposed to exploration at an even pace, users cluster exploration and exploitation into phases during their on-platform lifecycles[9]. These results suggest that existing theories of heterogeneous consumption need to account for fluctuating temporal dynamics in exploration. For example, established theories relying on instantaneous, one-time surveys (cf. Hargreaves and Bonneville-Roussy 2018) may capture listeners' preferences during peaks or troughs of musical exploration and miss longer-term trends.

Comparing Exploration and Diversity

Our work thus far measures *how much* people explore individual pieces of content (cf. Benson, Kumar, and Tomkins 2016; Kapoor et al. 2015), but does not explicitly consider *what* content is consumed nor the similarity between types of content. How do content-agnostic metrics like exploration compare to content-aware measures? One way of quantifying similarity is by using Spotify's track embedding described in the Data section, in which songs are placed depending on how often they are listened together.

Consider the *generalist-specialist* score ("GS-score"; see Waller and Anderson 2019). Given the tracks a user listens to, it measures how closely the corresponding track vectors are located together in the embedding space. So, users with higher scores are less diverse because their listening patterns are concentrated on similar tracks in a narrower space[10]. Compared to exploration, diversity defined this way quantifies content similarity but is also coarse-grained, whereas exploration is agnostic to similarity but granular – it explicitly measures individual pieces of content. Thus, do users who explore more also have more diverse musical interests?

We calculate GS-scores for each user's tracks across their entire trace and in each of their time windows. This latter metric gives us a coarse-grained measure of diversity over time. Suppose a user u listens to tracks $t \in T_i$ in the ith window, and each track is streamed $w_{i,t}$ times. Let $\vec{t}$ be each track's embedded vector. Then the user's ith centroid is:

$$\vec{\mu_i} = \frac{1}{\sum_{t \in T_i} w_{i,t}} \cdot \sum_{t \in T_i} w_{i,t}\vec{t}.$$

The user's GS-score in the ith window is their centroid's average cosine similarity with the content they consume, i.e.:

$$GS(\mu_i) = \frac{1}{\sum_{t \in T_i} w_{i,t}} \cdot \sum_{t \in T_i} w_{i,t} \frac{\vec{t} \cdot \vec{\mu_i}}{||\vec{t}|| \cdot ||\vec{\mu_i}||}$$

[9]Note that there are further between-group differences in discovery and turnover IQRs evident in Figures 7 (a) and 8 (a). We present a short follow-up analysis in the Appendix D.

[10]The GS-score is associated with other content-aware diversity metrics, such as entropy over track genres (Anderson et al. 2020a). The correlation is $r = -0.62$ in our dataset, suggesting that other content-aware metrics may yield comparable results. We also repeat the analyses in Figure 9 using genre entropy and found similarly faint relationships between exploration and entropy. We leave a full comparison of other content-agnostic and content-aware metrics to future work.

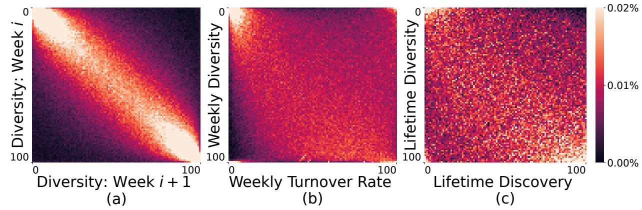

Figure 9: Distribution over diversity in adjacent weeks (a); diversity and turnover in the same week (b); diversity and discovery per stream across the entire trace (c). Variables are adjusted for activity and binned into percentiles.

The GS-score is known to have a slight correlation with activity. To correct for this, we use an *activity-adjusted* version. The adjusted GS-score for a user u during window i is the percentile rank of $GS(u,i)$ relative to $GS(v,j)$ for each other user v and window j that are in the same activity bin during window i. These bins are defined using the rounded square-root of stream count in each window. The adjusted diversity score is the adjusted GS-score with reversed ranks.

Figure 9 (a) illustrates the temporal consistency of adjusted diversity. For 10 sampled windows per user, it arranges each (u,i) pair on a heatmap according to the adjusted values of $GS(u,i)$ and $GS(u,i+1)$. We find the distribution of short-term, weekly diversity to be qualitatively almost identical to the distribution of long-term, yearly adjusted diversity presented by Anderson et al. 2020a. Thus, when we treat users as separate individuals each week, their diversity is still as temporally stable as found in users on aggregate. This supports the internal validity of our results.

But are diversity and exploration related? One may intuit that users occupying multiple, spread-out regions of the embedding space should explore more. We investigate this in Figure 9 (b), which visualizes the distribution of users according to their weekly diversity and turnover rate, both activity-adjusted for comparability. Similarly, Figure 9 (c) visualizes the distribution of adjusted diversity and discovery across the entire trace.

We find that, counter-intuitively, exploration and diversity appear to be unrelated. For example, the relationship between diversity and turnover is faint in comparison to the pattern seen in Figure 9 (a). Indeed, we only find a slight correlation of $r = 0.15$ between turnover and diversity on the weekly level for all users. Furthermore, diversity and global discovery are also nonequivalent: Figure 9 (c) reveals no clear association between the two metrics. The correlation between the two are $r = 0.19$, and $r = 0.17$ for the adjusted variant. Therefore, global exploration and lifetime diversity also appear to measure different quantities.

To highlight the distinction between exploration and diversity further, Figure 10 depicts weekly diversity over time. Notice that the rank-ordering of the age groups are the opposite of both global (Figures 4 and 5) and local exploration (Figure 6), especially in later years. The youngest group begins as specialists and diversifies over time, whereas listeners over the age of 45 became the least diverse of our cohort. These results are consistent with between-age variation

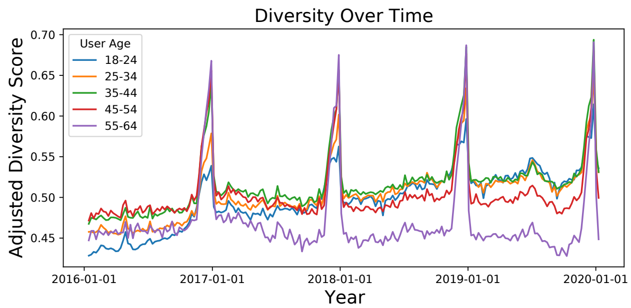

Figure 10: Expected weekly diversity score over time.

in diversity found in existing work (Anderson et al. 2020a). Thus, in context of listeners at different stages of their lifecycles, exploration and diversity are nonequivalent on Spotify.

Nonetheless, exploration and diversity still have some commonalities. Like our previous longitudinal analyses of exploration, diversity also follows similar annual cyclic patterns – it peaks significantly around the December holiday period relative to the rest of the year. This is consistent again with users shifting their listening habits according to seasonality[11]. Note, however, that the spikes in Figure 10 far outsize those in Figures 5, and 6. Together, these patterns indicate that *what* users explore changes (quantified through diversity) drastically more than *how much* they explore (quantified through discovery and turnover) during Christmas.

Another commonality is that younger listeners seem less affected by the holidays than older users, during which their traces have gentler diversity spikes than those found in the older groups. This too is consistent with the way in which seasonality affects younger users less in Figures 5 (a) and 6. Therefore, given that exploration and diversity have striking parallels in some aspects, it is even more surprising that they are generally nonequivalent. Younger users are exploiter-generalists, while older users are explorer-specialists.

Diversity and exploration thus paint pictures of different listener archetypes. On the one hand, a specialist explorer may frequently move between different pieces of content in a small region of the musical landscape. This could be, for example, someone who listens exclusively to movie soundtracks while both searching for older films and following new film releases. On the other hand, a generalist exploiter may listen to the same diverse set of content despite rarely searching for novelty. Consider a user who listens almost exclusively to a self-curated playlist of underground hardcore, K-pop, and bossa nova. They may be sufficiently satiated such that they do not crave exploration.

Discussion

Our work presents a large-scale, longitudinal analysis of how users on Spotify explore music over their off- and on-platform lifecycles. With respect to our research question, we find clear, pervasive differences between users of different ages in how they explore. Older listeners are consistently more likely to explore individual pieces of music at both the global and local level, whereas younger listeners are more likely to exploit content they are familiar with. These differences persist across different types of content and over the entire 4 year trace we studied. However, these trends are reversed for content diversity, with older listeners consuming less diverse content than younger listeners. Thus, in context of the multitude of existing theories addressing how variety-seeking behaviors change over time, our results show that older consumers explore narrower corpora of content.

In addition to its dependence on the off-platform the aging process, exploration also depends on user dynamics during the on-platform lifecycle. Between different time slices, users' exploratory patterns fluctuate according to seasonal cycles. Within individual users, exploration occurs in phases such that weeks with more exploration are likely to lead to successive exploration-heavy weeks. These observations emphasize the role of temporal dynamics in understanding variety-seeking behaviors. Thus, more longitudinal work on heterogeneous consumption is justified to better understand how platforms can facilitate exploration over time. For example, studies have modeled the evolution of behaviors over time (McAuley and Leskovec 2013; Moore et al. 2013; Benson, Kumar, and Tomkins 2016) and heterogeneity in narrow time slices (Anderson et al. 2020a; Schafer, Konstan, and Riedl 2002; Schedl and Hauger 2015), but not the evolution of heterogeneity on online platforms.

Insofar as exploratory behaviors are sensitive to the on- and off-platform lifecyle, we also find that they capture a different dimension of heterogeneous consumption that is not described by existing diversity metrics. At the atomic level of individual pieces of content, older users consistently *explore more* on both the long and short term. Thus, aging does not appear to focus users' attention on fewer pieces of content due to shrinking time horizons (Carstensen, Fung, and Charles 2003). However, older users are *less diverse*, which is concordant with the theory that preferences "set in" over time (Roberts, Walton, and Viechtbauer 2006; Srivastava et al. 2003). These results show that different measures of heterogeneity reveal the different ways in which listeners consume a variety of content. On the one hand, older users keep their listening interesting by consuming many distinct pieces of content; on the other, younger people keep their listening interesting through diverse music.

These findings motivate several prescriptive paths for helping online platforms and recommender systems facilitate variety-seeking behaviors. The differences in how users explore at various stages of their lifecycle, both on and off the platform, suggest that incorporating lifecycle information may improve platform design. For example, recommender systems could be augmented with exploration metrics to measure whether users are generally exploratory, and whether they are in an exploratory phase, and tailor recommendations accordingly. Indeed, recent work on ad timing indicates that this sort of temporal information can significantly impact receptiveness to recommendations (Saha et al. 2021). Furthermore, previous work has demonstrated that recommendations guided by heterogeneity are more likely to be satisfactory (Zhang et al. 2012; Schedl and Hauger

[11] See Appendix A for genre validation.

2015) and lead to positive retention outcomes (Anderson et al. 2020a), although there are trade-offs in engagement made when increasing consumption diversity (Holtz et al. 2020). And yet, our results illustrate that different measures of heterogeneity are not equivalent in general. Future work is therefore needed to evaluate how metrics like exploration and diversity can be combined to capture the ways in which users' needs for variety are satisfied on online platforms.

Appendix

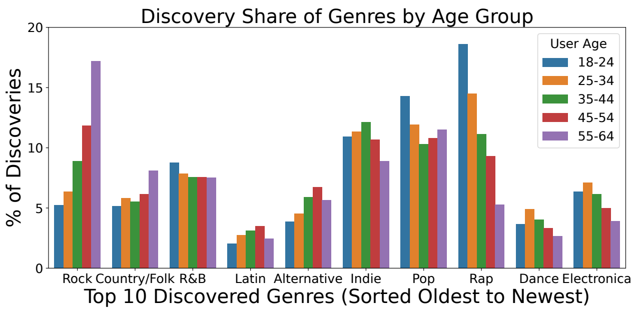

Figure 11: Discovery percentages of the 10 most discovered genres, sorted from left to right by how old the genres are.

A: Genres and exploration. Our study of exploration focuses on individual, atomic pieces of music without considering higher-level social constructs like genres. To what extent does analyzing genres aid in understanding our results? We consider discoveries of genre-specific tracks using Gracenote (cf. McKay and Fujinaga 2006) at the lowest level and present two such analyses to supplement our findings.

Children's music. We note that the 35–44 age group ranks higher than the rest in Figures 2 and 10, and floats above the 45-54 group in Figure 6. This could be explained by users increasing their exploration of children's material at the age of parenthood, which is evident when considering discoveries strictly of children's music. At 1.0%, the 35–44 age group discovers children's music at more than twice the rate of the other groups (0.3%, 0.4%, 0.5%, 0.3% of discoveries in increasing order of age). Similar musical peculiarities with this age group have been shown in related work (Way et al. 2019). Note however the limited effect size; more work will need to be conducted into this group.

Christmas music. To validate the annual cyclic patterns of discovery, turnover, and diversity peaking during Christmas, we consider the exploration of festive music during December compared to November and January. We find that discovery of Christmas music spikes dramatically during December, in which 2.8%, 3.5%, 4.7%, 5.3%, 6.6% of discoveries were Christmas music for each group in order of ascending age. This is substantially higher than November (between 1.4% and 1.9% for all groups) and January (below 0.2% for all groups). Furthermore, note that this effect skews towards older users – younger users are less affected by holiday seasonality than older listeners. Together, these observations validate the patterns in Figures 4, 5, 6, and 10.

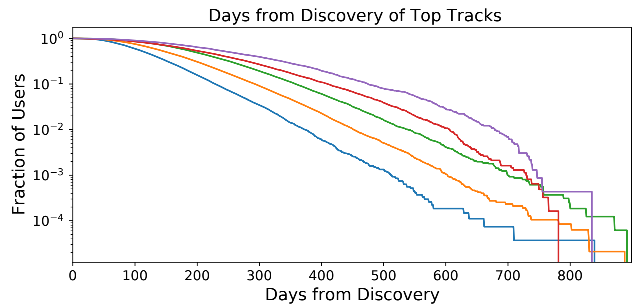

Figure 12: Time taken for a track to evolve from discovery by a user into one of the user's monthly top 10.

Novel content. To validate younger listeners' elevated discovery rates for newer music, we compare how much the 10 most-discovered genres are explored by different age groups in Figure 11, sorted by genre age. Genre ages are defined by the mean age of their tracks. Note that older genres are explored substantially more by older users, and newer genres more by younger users. Thus, there are significant exploration biases towards newer content by younger listeners. Indeed, Rock and Rap are respectively the most explored genres for the oldest and youngest groups, and simultaneously the oldest and fifth newest genres in our dataset. This reinforces our findings that exploration is contingent not only on the age of users but also on the age of content.

B: Time taken to convert discoveries. By analyzing c-discoveries in Figure 3, we find that the age differences in the prevalence of discoveries diminish as low values of c increases. However, different age groups still behave differently in terms of *how long* discoveries take to be converted.

For this analysis, we consider how a user's age is associated with their conversion of discoveries into their most-listened tracks. We analyze the middle two years of our trace (2017-9) and move a sliding 28-day window over each user's trace. For each window, we compute each user's personal top 10 by total stream count, i.e. the 10 highest c values that tracks can attain for each user in a 4-week period. Across tracks and users per age group, we take the mean time between a track's discovery and its entering the user's top 10.

Figure 12 depicts the complementary cumulative distributions of this metric. We find clear differences between all ages. The most repeatedly-streamed music for young users tend to be discovered more recently. 90% of the youngest group's top tracks were heavily streamed at most 230 days after discovery on average, whereas the same statistic for the oldest group is 475 days. The remaining groups' top tracks had means respectively at 294, 359, and 411 days. One interpretation is that younger users convert discoveries more quickly into their favorite tracks. This is consistent with our findings that young listeners are more likely to exploit already-discovered tracks, which may lead these tracks to be more easily "stuck" as favorites. Another interpretation is that older users have wider-spanning memories from which they can draw on earlier discoveries. This too is plausible given their higher number of unique tracks per stream.

The gaps between age groups are even more pronounced with respect to the top tracks' relative age, i.e. the time elapsed between release and entry into a user's top 10. In Figure 12, 90% of the tracks for each group were released at most 9, 11, 15, 19, and 24 years before being heavily streamed by individual users. Track ages were doubled for the oldest users compared to the youngest. Similarly, 61% of top tracks for the oldest listeners were released more than 15 years ago, compared to 8% for the youngest listeners.

To improve robustness, we also measured the mean time between the first listen and the cth listen of tracks for each group. We find the same ordering of age groups as Figure 12 for $c \in \{2, 3, 10, 20\}$. Thus, despite age differences in *conversion quantity* flipping for higher c shown in Figure 3, we still find that young users have lower *conversion times*.

C: Robustness to on-platform tenures. We note that a user's age may be related to how long they have been active on the platform, which could in turn impact our present analysis. For example, users who have used the platform for longer may be less prone to explore because they have already satiated their need to discover. However, we find that the correlation between users' join years and aggregate discovery rates is $r = -0.28$. This indicates that users with a longer history of engagement are more likely to explore.

Because younger users may have a shorter tenure on the platform, one may therefore question whether users' length of time on Spotify confounds the between-age effect presented here. To inspect this possibility, we identified the distribution of age groups over join dates. Only the youngest group was skewed towards recent join years (up to 2015, the year before the trace begins). The modal years for the remaining age groups were 2012 for those between 25-34, 2011 for those between 35-44, and 2014 for those over 45. Thus, age is actually poorly correlated with join date.

To further ensure that tenure length do not confound the observations above, we repeated the same analysis in Figures 4 and 5 while conditioning by registration year. We find that the temporal patterns and age group differences are virtually identical in each iteration of this figure. We also measured discoveries over time split by registration year while conditioning on age. Again, we find that later joiners explored consistently less across all age groups. Thus, these robustness checks indicate that registration dates are not a major confound. Instead, earlier registration and older age are both associated with exploration in distinct ways.

D: Age-dependent differences in exploration phases. One observation worth noting is the subtle differences between age groups in Figure 7 (a) and Figure 8 (a). This gives rise to a further question: are age-dependent differences in phases more prevalent in discoveries, or are they more prevalent in turnover? To more closely evaluate this, we analyze the delta in weekly exploration IQRs between users and their permuted selves. In other words, this is a measure of how much more skewed users are in reality than their behaviors under the null hypothesis that exploration is evenly dispersed.

We find the discovery IQR deltas for each group to be 0.035, 0.048, 0.064, 0.069, 0.076 ($p < 0.001$ for all ages using one-tailed Holm-corrected Mann-Whitney U tests on adjacent age groups), indicating that older users are more likely to group discoveries into phases than younger users. In comparison, we find the turnover IQR deltas for each group to be 0.114, 0.112, 0.118, 0.121, and 0.118 ($p < 0.05$ only for 25-34 vs 35-44 and 35-44 vs 45-54).

Thus, temporal phases are found both in discovery and turnover, but the former has substantially more between-group variance. This corroborates our findings that young users have lower cumulative discovery rates (Figure 2). Discoveries must draw from a user's unique set of tracks across their lifetime, so a lower weekly discovery ceiling would lead to lower weekly IQRs. In contrast, turnover only depends on the preceding week and is not subject to this effect.

References

Anderson, A.; Maystre, L.; Anderson, I.; Mehrotra, R.; and Lalmas, M. 2020a. Algorithmic effects on the diversity of consumption on spotify. In *Proceedings of The Web Conference*, 2155–2165.

Anderson, I.; Gil, S.; Gibson, C.; Wolf, S.; Shapiro, W.; Semerci, O.; and Greenberg, D. M. 2020b. "Just the Way You Are": Linking Music Listening on Spotify and Personality. *Social Psychological and Personality Science*.

Batcho, K. I. 1995. Nostalgia: A psychological perspective. *Perceptual and Motor Skills*, 80(1): 131–143.

Benson, A. R.; Kumar, R.; and Tomkins, A. 2016. Modeling user consumption sequences. In *Proceedings of the International Conference on World Wide Web*, 519–529.

Bourdieu, P. 1984. *Distinction: A social critique of the judgement of taste*. Harvard University Press.

Carstensen, L. L.; Fung, H. H.; and Charles, S. T. 2003. Socioemotional Selectivity Theory and the Regulation of Emotion in the Second Half of Life. *Motivation and Emotion*.

Costa Jr, P. T.; McCrae, R. R.; Zonderman, A. B.; Barbano, H. E.; Lebowitz, B.; and Larson, D. M. 1986. Cross-sectional studies of personality in a national sample: II. Stability in neuroticism, extraversion, and openness. *Psychology and aging*, 1(2): 144.

Danescu-Niculescu-Mizil, C.; West, R.; Jurafsky, D.; Leskovec, J.; and Potts, C. 2013. No country for old members: User lifecycle and linguistic change in online communities. In *Proceedings of the International Conference on World Wide Web*, 307–318.

Datta, H.; Knox, G.; and Bronnenberg, B. J. 2018. Changing their tune: How consumers' adoption of online streaming affects music consumption and discovery. *Marketing Science*, 37(1): 5–21.

Ferwerda, B.; Tkalcic, M.; and Schedl, M. 2017. Personality traits and music genre preferences: how music taste varies over age groups. In *Workshop on Temporal Reasoning in Recommender Systems (RecTemp) at the ACM Conference on Recommender Systems*, volume 1922, 16–20.

Gibson, E. J. 1988. Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Annual Review of Psychology*, 39(1): 1–42.

Hargreaves, D. J. 1982. The development of aesthetic reactions to music. *Psychology of Music*.

Hargreaves, D. J.; and Bonneville-Roussy, A. 2018. What is 'open-earedness', and how can it be measured? *Musicae Scientiae*, 22(2): 161–174.

Holbrook, M. B. 1993. Nostalgia and consumption preferences: Some emerging patterns of consumer tastes. *Journal of Consumer Research*, 20(2): 245–256.

Holtz, D.; Carterette, B.; Chandar, P.; Nazari, Z.; Cramer, H.; and Aral, S. 2020. The Engagement-Diversity Connection: Evidence from a Field Experiment on Spotify. In *Proceedings of the 21st ACM Conference on Economics and Computation*, 75–76.

Kapoor, K.; Kumar, V.; Terveen, L.; Konstan, J. A.; and Schrater, P. 2015. "I like to explore sometimes": Adapting to Dynamic User Novelty Preferences. In *Proceedings of the ACM Conference on Recommender Systems*, 19–26.

Kidd, C.; and Hayden, B. Y. 2015. The Psychology and Neuroscience of Curiosity. *Neuron*, 88(3): 449–460.

LeBlanc, A.; Sims, W. L.; Siivola, C.; and Obert, M. 1996. Music style preferences of different age listeners. *Journal of Research in Music Education*, 44(1): 49–59.

McAuley, J. J.; and Leskovec, J. 2013. From amateurs to connoisseurs: modeling the evolution of user expertise through online reviews. In *Proceedings of the International Conference on World Wide Web*, 897–908.

McKay, C.; and Fujinaga, I. 2006. Musical genre classification: Is it worth pursuing and how can it be improved? In *ISMIR*, 101–106. Citeseer.

Mikolov, T.; Sutskever, I.; Chen, K.; Corrado, G. S.; and Dean, J. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems*, 3111–3119.

Moore, J. L.; Chen, S.; Turnbull, D.; and Joachims, T. 2013. Taste Over Time: The Temporal Dynamics of User Preferences. In *Proceedings of the International Society for Music Information Retrieval Conference*, 401–406.

Nicklaus, S.; Boggio, V.; Chabanet, C.; and Issanchou, S. 2005. A prospective study of food variety seeking in childhood, adolescence and early adult life. *Appetite*, 44(3): 289.

Novak, D. L.; and Mather, M. 2007. Aging and variety seeking. *Psychology and Aging*, 22(4): 728.

Park, M.; Thom, J.; Mennicken, S.; Cramer, H.; and Macy, M. 2019. Global music streaming data reveal diurnal and seasonal patterns of affective preference. *Nature Human Behaviour*, 3(3): 230–236.

Park, M.; Weber, I.; Naaman, M.; and Vieweg, S. 2015. Understanding Musical Diversity via Online Social Media. In *Proceedings of the International Conference on Web and Social Media*.

Peterson, R. A. 2005. Problems in comparative research: The example of omnivorousness. *Poetics*, 33(5-6): 257–282.

Ratner, R. K.; Kahn, B. E.; and Kahneman, D. 1999. Choosing less-preferred experiences for the sake of variety. *Journal of Consumer Research*, 26(1): 1–15.

Roberts, B. W.; Walton, K. E.; and Viechtbauer, W. 2006. Patterns of mean-level change in personality traits across the life course: a meta-analysis of longitudinal studies. *Psychological Bulletin*, 132(1): 1.

Saarikallio, S. 2011. Music as emotional self-regulation throughout adulthood. *Psychology of Music*, 39(3): 307–327.

Saarikallio, S.; and Erkkilä, J. 2007. The role of music in adolescents' mood regulation. *Psychology of Music*.

Saha, K.; Liu, Y.; Vincent, N.; Chowdhury, F. A.; Neves, L.; Shah, N.; and Bos, M. W. 2021. AdverTiming Matters: Examining User Ad Consumption for Effective Ad Allocations on Social Media. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–18.

Schafer, J. B.; Konstan, J. A.; and Riedl, J. 2002. Meta-recommendation systems: user-controlled integration of diverse recommendations. In *The International Conference on Information and Knowledge Management*, 43–51.

Schedl, M.; and Hauger, D. 2015. Tailoring music recommendations to users by considering diversity, mainstreaminess, and novelty. In *Proceedings of the Conference on Research and Development in Information Retrieval*, 947–950.

Soto, C. J.; John, O. P.; Gosling, S. D.; and Potter, J. 2011. Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology*, 100(2): 330.

Srivastava, S.; John, O. P.; Gosling, S. D.; and Potter, J. 2003. Development of personality in early and middle adulthood: Set like plaster or persistent change? *Journal of Personality and Social Psychology*, 84(5): 1041.

Tarrant, M.; North, A. C.; and Hargreaves, D. J. 2002. Youth identity and music. *Musical Identities*, 13: 134–150.

Tervo, J. 2001. Music therapy for adolescents. *Clinical Child Psychology and Psychiatry*, 6(1): 79–91.

Vodanovich, S.; and Kass, S. 1990. Age and gender differences in boredom proneness. *Journal of Social Behavior and Personality*, 5: 297–307.

Waller, I.; and Anderson, A. 2019. Generalists and specialists: Using community embeddings to quantify activity diversity in online platforms. In *Proceedings of the International Conference on World Wide Web*, 1954–1964.

Way, S. F.; Gil, S.; Anderson, I.; and Clauset, A. 2019. Environmental Changes and the Dynamics of Musical Identity. In *Proceedings of the International Conference on Web and Social Media*, volume 13, 527–536.

Zhang, M.; and Hurley, N. 2008. Avoiding monotony: improving the diversity of recommendation lists. In *Proceedings of the Conference on Recommender Systems*, 123–130.

Zhang, Y. C.; Séaghdha, D. Ó.; Quercia, D.; and Jambor, T. 2012. Auralist: introducing serendipity into music recommendation. In *ICWSDM*, 13–22.

Zuckerman, M.; and Neeb, M. 1979. Sensation Seeking and Psychopathology. *Psychiatry Research*, 1(3): 255–264.