Tracing the mechanisms of cultural diversity through 2.5 million individuals' music listening patterns

Harin Lee[1, 2, 3, 4, ⁎], Nori Jacoby[3], Romain Hennequin[1, +], Manuel Moussallam[1, +]

[1] Deezer Research, Paris, France
[2] Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[3] Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany
[4] Department of Life Sciences, Leipzig University, Leipzig, Germany

⁎ hlee@cbs.mpg.de
+ research@deezer.com (equal contribution)

Keywords: *cultural diversity, population size, group diversity, collective action, music streaming*

Large cities are culturally diverse. Does this diversity stem from a mix of people with varied demographic backgrounds, or from each individual seeking a broader spectrum of cultural experiences? Due to the lack of data on individual behavioural patterns and demographics, previous research could not discern individuals' contributions to collective cultural outcomes. Leveraging data from over 2.5 million French, Brazilian, and German listeners, comprising 250 million real-world listening behaviours, we trace the mechanisms contributing to cultural diversity across regional populations within these countries. We assess the collective shared musical repertoire in each region, while concurrently measuring each individual's breadth of music engagement through their personal listening histories. Our analysis reveals that an increase in diversity at the collective scale is associated with the size of its population, aligning with previous research on cultural evolution and urban scaling law. We extend this understanding by demonstrating that an individual's cultural breadth also expands with population size, consistently across all three countries. Although demographic factors such as age, gender, income, immigration, education, and social connections mediate these trends, particularly in the most populated areas, they do not fully explain them. This potentially suggests that large cities are culturally diverse not only because they bring together people from varied backgrounds, but also due to the greater opportunities for cultural interaction and exchange that urban settings facilitate.

Introduction

Urban life is vastly different from the experience in smaller cities and towns. Inhabitants of large metropolitan areas tend to live a faster pace lifestyle[1,2], engage in more social interactions[3], and the limited physical spaces allow for more rapid dissemination of information but also of diseases[4,5]. Recent census data from various countries reveal a consistent pattern. As cities grow in size, their collective outputs in various aspects, such as economy and education, surpass linear growth expectations[6–9], a phenomenon known as the urban scaling law. Cultural richness and complexity have also been shown to expand with population size. Computer simulations[10–15] and experimental studies[16–19] have demonstrated that larger populations are not only effective at preserving cultural complexity and diversity but can also stimulate innovation through increased cultural exchanges. Despite these insights, the exact mechanisms and factors contributing to cultural diversity remain a topic of debate[20–23].

Of all the possible mechanisms, two hypotheses offer concrete predictions. The 'demographic mixing' hypothesis suggests that cultural diversity in densely populated cities is a result of demographic variances[1,24–26], attracting individuals from diverse backgrounds, as seen in megapolis like Paris, Berlin, and São Paulo. In contrast, the second hypothesis rather focuses on the wider 'cultural breadth' of individuals living in those cities, proposing that continuous exposure to diverse and rich cultural inputs impacts one's psychology and broadens their cultural taste[27–30], leading to a communal shift towards greater diversity.

Previous studies have struggled to distinguish between these hypotheses, primarily due to the reliance on aggregated population-level data that overlooks individual behavioural patterns and demographics. To bridge this gap, we analyse a very large dataset of music-listening events in the real world. Music serves as an ideal test-bed for examining questions about cultural diversity, with its universal prevalence [31,32] and cultural distinctiveness[33–35], and it being a significant part of people's everyday lives[36,37]. Prior research, such as a study on Irish folk music by Street et al.[38] showed that tunes played by a larger community of musicians evolve into more varied versions and exhibit increased melodic complexity. However, no study has yet undertaken a large-scale and cross-cultural analysis of the influence of size and individual demographics on both collective and individual engagements of culture. This in part is due to the challenges faced in obtaining extensive individual level data.

Over four weeks, we tracked over 250 million listening events across 2,544,548 users on Deezer, a global music streaming platform available in 187 countries featuring over 90 million tracks. We collected data from France, Brazil, and Germany, which are the three countries with the largest percentage of Deezer users relative to their general populations. Using municipalities and NUTS unit geographic boundaries, we assessed population-wide diversity in listening patterns by examining the alignment of music listening between individuals. Concurrently, we examined these individuals' breadth of music engagement by tracing their listening histories. This dual perspective of diversity allows us to decipher different ways in which size affects cultural diversity outcomes of between and within individuals. Moreover, the cross-cultural aspect of the data allows us to test the generalisability of the phenomenon.

To assess the extent to which level of diversity might arise from the heterogeneity of the population (i.e., demographic mixing) or the broader spectrum of cultural engagement (i.e., cultural breadth), we account for various socio-demographic factors using simple linear regressions and causal inference tools. The demographic factors include self-reported age and gender, as well as inferred levels of income, education, immigrant status, access to music venues, and social connections. The inferred data is derived from various sources, offering detailed granularity at the level of over 35,000 municipalities, equivalent to postcode area resolution. We achieve this by pinpointing each user's home location and matching it to the corresponding demographic information from these sources. This expansive, spatially and temporally detailed dataset (openly available at https://github.com/harin-git/mus-div) offers us a unique opportunity to understand how local population size and demographic nuances influence multiple layers of cultural diversity.

Results

Fig. 1a shows the user distribution in our data. The data spanned over four weeks in March 2023 across users in France, Brazil, and Germany. After filtering for users only with reliable geolocations (User sampling in Methods), the final dataset encompassed over 2.5 million users (Fig. 1a) across France (N = 1,506,899), Brazil (N = 816,101), and Germany (N = 221,548). Detailed demographics and comparison with census data can be seen in the 'User demography' section in Methods. To ensure a balanced contribution from each user, we drew a random 100 unique streams per user. Users were grouped according to the geographical boundaries of regional units (User sampling in Methods), and we discarded regions with less than 200 unique users to deal with noisy data, and to ensure further anonymisation at the aggregate level. We analysed the data after anonymisation and the computations were not used to derive any commercial user profiling of any kind. All analysis scripts used in the study and aggregated anonymized data are openly available (Data availability).

More distinct music interests in large metropolitan areas

We begin by assessing the diversity of music listening between individuals, analysing how similar or dissimilar one's listening is to another residing in the same area. As a measure for between-individual diversity (BID), we adopt Hill's number, a widely used method for measuring cultural diversity and allows for standardised comparisons[39–41] (Measuring diversity in Methods). A high overlap of songs among individuals indicates a more shared consumption of musical repertoire (i.e., low diversity), while a low overlap indicates there being a greater distinction between individuals (i.e., high diversity). To ensure that varying user sample sizes across regions do not skew our diversity estimates, this analysis and all statistics reported in the paper are standardised by drawing an equal number of samples at every level of analysis and computing bootstrap estimates (Measuring diversity and Statistical analysis in Methods).

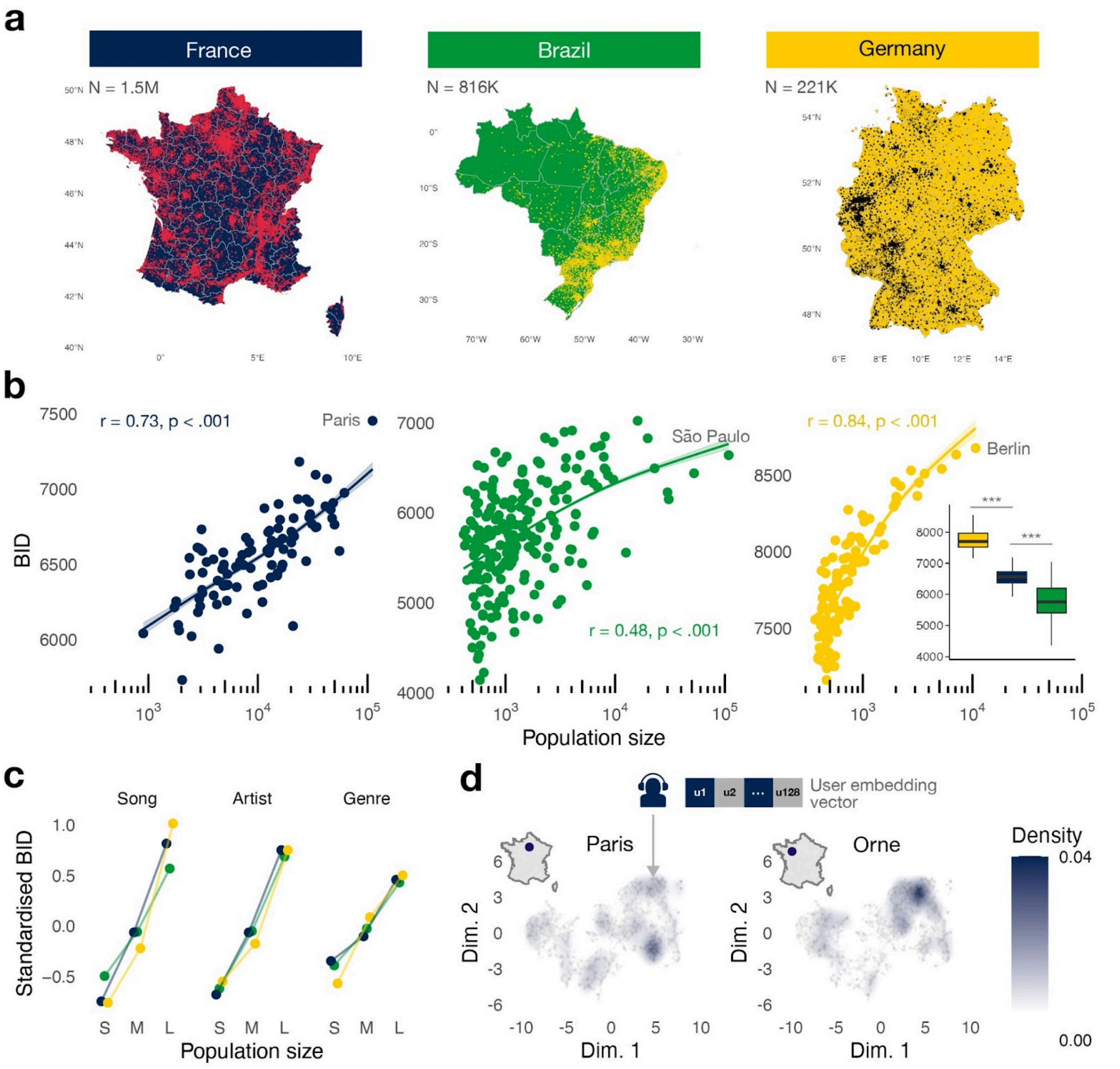

Fig. 1: Individuals' music listening is more distinct from one another in large metropolitan areas, resulting as dispersed pockets of diverse music interests.

(a) Geolocations of over 2.5 million users sampled from France, Brazil, and Germany. **(b)** Between-individual diversity (BID) is measured using Hill's number and computed across the geographical boundaries of NUTS3 units in France and Germany, and municipalities of Brazil (Measuring diversity in Methods). Each dot represents the mean estimate of sampling 10,000 unique music streams across bootstrap simulations

(Statistical analysis in Methods). Curved lines are generalised additive model (GAM) fitting between BID and population size of log base 10. Shaded areas and error bars represent 95% CI. The inset boxplot shows cross-country differences using the post-hoc Tukey test with significance levels at $p < .001$ (***), $p < .01$ (**), and $p < .05$ (*) after correcting for multiple comparisons. **(c)** Alternative units of analysis, grouping streams by 'artist' and 'musical genre', with standardised units and three quantile population size categories. **(d)** Measuring the dispersion of regional users in the user embedding space as an alternative method for assessing BID. Each user is represented as a 128-dimensional vector defined by their centroid, which is computed from all music they have listened to in the past 28 days (User embedding in Methods). Dimensionality reduction is performed using UMAP by balancely sampling users across regions. Users that are closer in this space tend to have more similar music interests. Using two regions as examples, kernel density estimation is applied to indicate the dispersion of regional users in the general space (Dispersion in user embedding in Methods). More dispersed density is observed across Parisians, indicating they have more distinct listening patterns from one another. In contrast, users from Orne, a small region with less than 300,000 inhabitants, demonstrate a more concentrated hotspot, suggesting a more homogenous taste among its residents.

Using geographical boundaries of NUTS3 units in France and Germany, and municipalities of Brazil (User demographic in Methods), we assess the level of BID relative to the region's population size (derived from the number of users in the region and log-transformed with base 10). In all three countries, there were strong positive correlations (Fig. 1b; France: Pearson $r = 0.73$, 95% CI = [0.70, 0.76]; Brazil: $r = 0.48$ [0.46, 0.49]; Germany: $r = 0.84$ [0.83, 0.86]; see Supplementary Table 1 for additional correlations). This indicates that individuals living in urban areas tend to have more distinct listening patterns from one another than those in rural areas. On average, large metropolitan areas with more than 500,000 inhabitants had 4.7% [4.3, 5.1] higher BID compared to small areas with less than 100,000 inhabitants (Cohen's $d = 1.21$ [1.10, 1.33], $p < .001$; Statistical analysis in Methods). Interestingly, our analysis also revealed notable cross-cultural differences (Fig. 1b inset; one-way ANOVA: $F(2, 424) = 671$, $p < .001$). BID in Germany (Mean = 7,77 [7,71, 7,84]) was substantially higher than in France (Mean = 6,55 [6,49, 6,60], $d = 3.83$, $p < .001$), and in Brazil (Mean = 5,76 [5,68, 5,84], $d = 3.89$, $p < .001$). Together, these results suggest a country-specific characteristic in the amount of shared music repertoires, alongside a country-independent, universally positive trend between BID and local population size.

To examine the robustness of our results, we conduct a series of control experiments. First, we expand our groupings of analysis and test beyond frequencies of 'songs', to 'artists' and 'musical genres' (Fig. 1c). Second, we account for various biases that might stem from our user filtering criteria (Supplementary Fig. 2) and algorithmically recommended streams (Supplementary Fig. 4). Third, we explore alternative population size metrics, using census population and density (Supplementary Fig. 5), and employ alternative diversity measures (Supplementary Table 2). We observed consistent patterns across all of these cases.

Finally, for further validation, we adopt an alternative approach for measuring BID, recognising that measuring diversity purely by frequency counts might overlook the relationships between the songs. We leverage a high-quality user embedding space constructed from the general user behaviour on the platform (User embedding in Methods).

We then use this embedding space to assess the dispersion amount of regional users in each region, whereby more dispersion would indicate more heterogeneity in music interests. Using France as our case example, we subsample a balanced number of users across the NUTS3 unit regions. Each user is represented as a 128-dimensional user vector, defined by their centroid computed across all music they have listened to in the last 28 days (User embedding in Methods). In this space, two users with similar music interests are situated closer together (e.g., both are K-pop fans), or further apart if they have more distinguished interests (e.g., K-pop versus metal fans). Fig. 1d illustrates this space as a two-dimensional UMAP, as an example, comparing the distribution of the users from the capital city of France, Paris, with users from a small region in the west, Orne. Kernel density estimation (KDE) is then overlayed to capture the density distribution of regional users in the general space (Dispersion in user embedding in Methods). The two density distributions demonstrate that the Parisian listeners are more scattered and dispersed across a broader space, indicative of there being pockets of diverse music interests. Conversely, the listeners from Orne are more concentrated within a narrower area, demonstrating that they have more shared, homogenous tastes. The amount of dispersion in each region, measured as the variance in the embedding space, revealed a similar correlation with population size ($r = 0.48$ [0.31, 0.62], $p < .001$), and also strongly correlated with the frequency-based BID metric ($r = 0.76$ [0.66, 0.83], $p < .001$).

To summarise, our analyses consistently reveal significant disparities in listening patterns between individuals residing in large metropolitan areas and those in rural areas.

Wider personal breadth of music engagement in large metropolitan areas

While previous studies were limited to measuring diversity at the population scales (equivalent to our assessment of BID), we further leverage the individual-level data to understand how an individual's breadth of music engagement may be associated with the size of the region they reside in. To assess within-individual diversity (WID), we adopt a metric known as the Generalist-Specialist Score (GS-Score) of music engagement[42,43] by tracing each individual's past 28 days of listening history (Measuring diversity in Methods). The GS-Score calculation is similar to our approach for assessing regional user dispersion in the previous section, but with a key distinction: we now measure the dispersion across songs within individual users' listening habits, rather than the dispersion among users. We utilise high-dimensional song embeddings computed based on the co-occurrence of millions of songs from user-generated playlists and listening behaviour (Song embedding in Methods). The dispersion of the songs a user listens to in this space is then used as a proxy for their general breadth of engagement, and it forms the foundation for many recommendation systems. Should one be a specialist, their consumption will concentrate on a narrow area of focus (e.g., listens only to K-pop), or be a generalist and engage with more diverse content (e.g., listens to K-pop but also metal and classical).

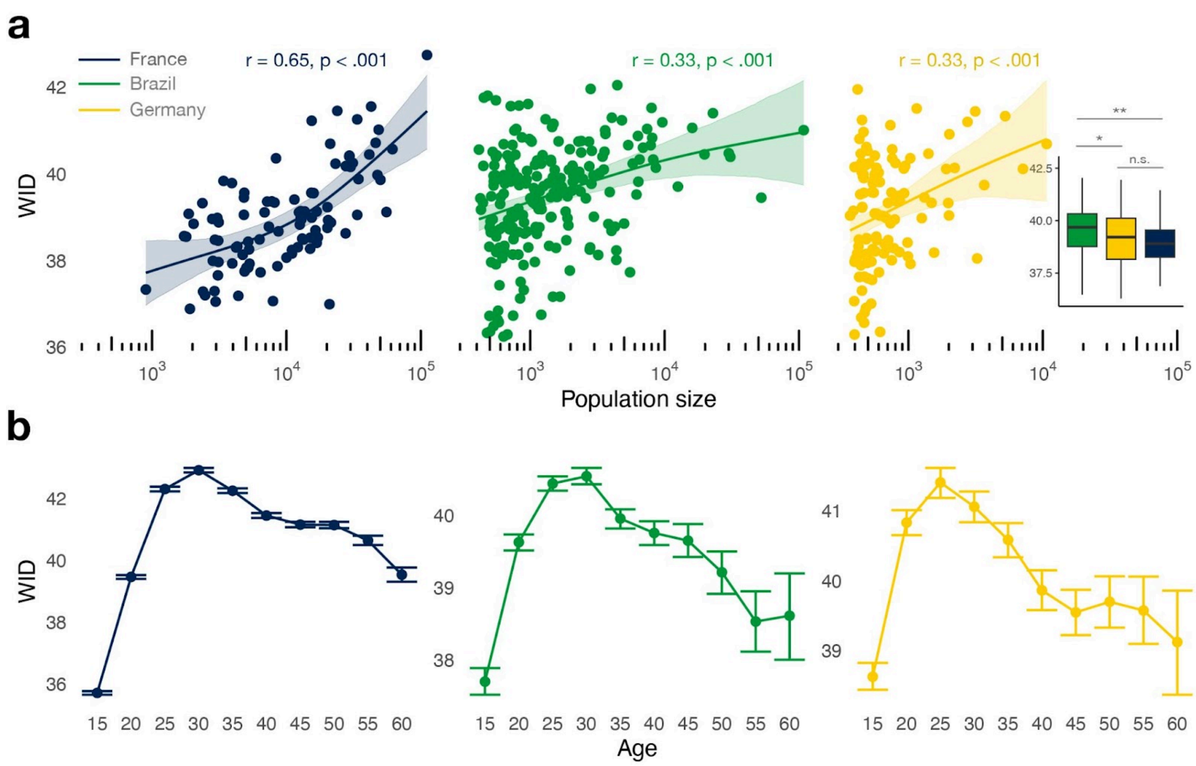

Fig. 2: Individuals living in large metropolitan areas engage with more diverse music. One's breadth of this engagement follows an inverted U-shape trajectory over the course of life.

(a) Relationship between population size and within-individual diversity (WID) with values ranging from 0 (least diverse) to 100 (most diverse). The metric captures how dispersed one's musical engagement is in the general music embedding space that defines relationships between songs (Measuring diversity in Methods). The inset boxplot shows cross-country differences using post-hoc Tukey test with significance levels at $p < .001$ (***), $p < .01$ (**), and $p < .05$ (*) correcting for multiple comparisons. **(b)** WID as a function of age. Cross-culturally, they form an inverted U-shape trajectory with the peak of most diverse exploration being around the late 20s. Each dot represents the mean estimate of sampling 100 random unique individuals across bootstrap simulations (Statistical analysis in Methods). Curved lines are GAM fitting with shaded area and error bars representing 95% CI.

Using the same sampled users and geographical units, we assess the level of WID relative to the region's population size. Mirroring the results of BID, we observed the same, but reduced, positive trends across all three countries (Fig. 2a; France: $r = 0.65$ [0.60, 0.68], $p < .001$; Brazil: $r = 0.33$ [0.30, 0.35], $p < .001$; Germany: $r = 0.33$ [0.31, 0.35], $p < .001$; see Supplementary Fig. 6 for correlation between WID and BID). Large areas (over 500,000) had 2.5% [1.3, 3.6] higher WID than smaller areas (less than 100,000; $d = 0.62$ [0.32, 0.93], $p < .001$).

While we previously saw substantial cross-cultural differences in the level of BID, the three countries exhibited similar levels of WID (Fig. 2a inset). Brazil (Mean = 39.5 [39.3, 39.7]) had the highest level of WID compared to Germany (Mean = 39.1 [38.9, 39.4]; $d = 0.30$, $p = .01$)

and France (Mean = 39.0 [38.7, 39.2]; $d = 0.43$, $p < .001$), but the effects were much smaller in size, and there was no significant difference between Germany and France ($p > .05$). These results suggest that individuals who reside in more populated areas tend to engage with a broader spectrum of music content, potentially stimulated by more diverse inputs coming from the richer cultural environment. At the same time, small differences between countries in the general size of this breadth indicate that an individual's cultural background does not play a substantial role in shaping this breadth.

Among many factors that may influence one's breadth of engagement, previous experimental research[44,45] and observational data of online behaviour[46] have noted 'age' as a strong predictor of consumption diversity — suggesting that people tend to explore and discover more during adolescence while their preferences become more stable with age. However, to our knowledge, no study to date has empirically demonstrated this trend in music consumption with detailed annual granularity across the lifespan.

As validation of our method for measuring within-individual diversity and to extend previous findings, we analyze WID as a function of age. Fig. 2b demonstrates a remarkably consistent, inverted U-shape pattern across all three countries. This cross-cultural pattern shows that individuals' musical experiences broaden from their teenage years into their early 20s, reach a peak in their late 20s, and then gradually narrow as they get older. Notably, this trend persisted even when accounting for variations in baseline streaming activities across different age groups, as younger individuals tend to stream more music (Supplementary Fig. 7).

Diversity is largely explained by demographic factors, but not entirely

The strong influence of age on WID sets a good example of the necessity to control for socio-demographic confounders that can bias our estimate of the relations between the two levels of diversity and population size. Larger cities tend to attract younger, more educated, and wealthier individuals who may demonstrate 'cultural omnivore' tendencies, embracing broader musical preferences from elite to popular genres[30,47–49]. Metropolitan areas are often also international hubs and where diverse cultural activities concentrate, such as there being more music events. Such unique properties of large cities are in stark contrast in their demographics compared to their rural counterparts. In fact, several past studies have found the influence of population size on diversity to reduce, or even diminish, when controlling for other demographic differences, migration, and social connections[48–52].

To address the question of whether the increased diversity we observe is a result of demographic differences between urban and rural areas, we go beyond correlational analyses. We focus on France, given the highest spatial granularity of our user sample and the availability of dense high-quality demographics. We collect comprehensive data on various socio-demographic attributes at the level of over 35,000 municipalities and 96 NUTS3 regional units (see Supplementary Fig. 8 for demographic disparities by regions' size). First, we gather regional census data from Eurostats and The National Institute of Statistics and Economic Studies (INSEE), covering median income, proportion of immigrants, and proportion of university degree holders (see Supplementary Fig. 9 for map visualisations). Second, we approximate the accessibility to music events in each region by mapping the locations of 6,618 music venues across the country using the SongKick database — a popular website for tracking music events (see Supplementary Fig. 10 for map visualisation). Third, we include data from Facebook to approximate the amount of international social connections. Lastly, we obtain self-reported age and gender from Deezer's registration information. All details regarding the justification of considering these demographic variables as potential confounders, along with the data source and the collection method, can be seen in the 'Socio-demographics' section in Methods.

We noticed significant correlations between demographic variables and diversity, as indicated by the correlation matrix and descriptive statistics provided in the supplementary information (Supplementary Fig. 11 and Table 3). To further investigate this relationship, we perform simple linear regressions to assess the extent to which demographic factors account for the diversity. We first categorise user samples into three population quantile categories: small (census inhabitant Median = 228,750), medium (Median = 542,122), and large (Median = 1,260,378). Next, for each population size category, we separately regress all demographic factors against BID and WID. The adjusted R-square values showed that demographic variables explain 27.4% of BID and 30.6% of WID in small regions, 52.2% of BID and 39.3% of WID in medium regions, and 80.1% of BID and 83.7% of WID in large regions. This suggests that high levels of BID and WID in the largest metropolises mainly arise from their demography.

Nevertheless, such a simple model ignores the interaction among the variables, and these interactions can bias the observed estimates in various ways (e.g., collider bias; see Cinelli, Forney & Pearl[55] on 'bad controls'). Thus, we introduce a more complex model by making assumptions about the relationships among these variables based on the existing literature (Socio-demographics in Methods). We illustrate this in the form of a Directed Acyclic Graph (DAG) and check that implied conditional independencies are met (Causal inference in Methods). DAG provides a clear and efficient method to identify, present, and hypothesise the causal relationships between variables[56,57]. Importantly, DAG allows to identify which confounder should be controlled or left uncontrolled, rather than controlling for every imaginable covariates (i.e., throwing everything into the sink). Using this DAG, we test the direct effect of population size on BID (Fig. 3a) and WID (Fig. 3b) after controlling for demographic variables.

In the two models we test, we first identify the confounders (red in figure; i.e., minimal adjustment sets[55]) and adjust them to obtain a direct estimate of exposure (yellow in figure) on the outcome (blue in figure). Next, we assign inverse propensity weights (IPW) to each user based on the average treatment overlapping (ATO) population method[58,59]. This procedure is to balance the distribution of the confounding variables across the groups. IPW is a widely used modern causal inference technique to create an artificial scenario that mimics a randomised trial (for a review on causal inference methods, see Chatton & Rohrer, 2023[59]). The standardised mean difference (SMD) tests (Supplementary Fig. 12) and the empirical cumulative distribution function (ECDF) of the weights (Supplementary Fig. 13) indicated that the groups are well balanced after applying the weights. Lastly, we ran bootstrap simulations of the models by sampling users and assigning new weights in each iteration to obtain confidence estimates (Statistical analysis in Methods).

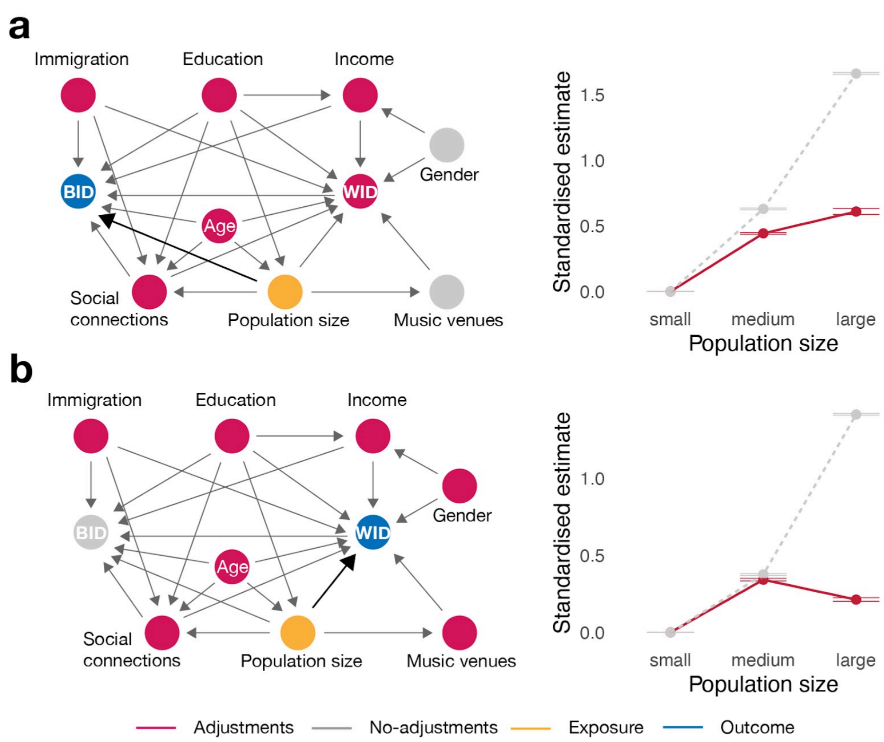

Fig. 3: Demographic variables largely account for the diversity in the most populous regions and reveal distinctions between BID and WID.

(a) BID and **(b)** WID outcomes after adjusting for demographic confounders. The models are illustrated in the form of DAG and inverse propensity weights are applied to users categorised into three quantile population groups. Adjusted variables in the models (i.e., minimal adjustment set) are coloured in red, whereas unadjusted variables are coloured in grey (Causal inference in Methods). Corresponding colour coding is used to present the model outcomes, comparing adjusted (solid, red) and unadjusted (dashed, grey) effects across the population size groups. The estimates are standardised using Z-scores with intercepts matched at the 'small' population group for effect size comparisons. Error bars represent 95% CI from bootstrapped model simulations (Statistical analysis in Methods).

Fig. 3a,b shows the results contrasting the unadjusted effects (dashed and grey lines) with those adjusted using IPW (red and solid lines). The variables are standardised using Z-scores and the intercepts are matched at the 'small' population group for effect size comparisons (see Supplementary Fig. 14 for trends stratified by each confounder separately). After adjusting for demographic differences, the general trend remained:

medium and large regions exhibited greater cultural diversity than smaller ones (ps < .001). However, it also revealed notable nuances, especially in the largest population quantiles. Consistent with the simple regression analysis reported earlier, demographic factors accounted for much of the cultural diversity in the largest regions. However, this was more pronounced for WID than BID (reduction in standardised estimate: WID = 1.20 [1.19, 1.21], BID = 1.05 [1.03, 1.08]). As a result, BID remained a consistent linear increase with population size (standardised estimate from intercept: medium = 0.44 [0.44, 0.45], large = 0.60 [0.59, 0.62]), whereas WID levelled off after the medium quantile (medium = 0.34 [0.34, 0.35], large = 0.21 [0.20, 0.22]).

To conclude, the findings indicate that while demographic mixing contributes significantly to the cultural diversity of highly populated areas (i.e., capitals), it does not entirely account for the ongoing increasing trend in diversity between individuals. Conversely, the plateau observed in WID beyond a certain population threshold (around 500,000 inhabitants) implies that there is a saturation point for individuals' cultural exploration.

Discussion

Prior observations of heightened cultural diversity in large populations were limited in understanding whether this increase is driven by the population's demographic heterogeneity, or by broader cultural engagements of individuals. Using music listening patterns in the real world, we sought to decipher this and study the mechanisms and factors contributing to cultural diversity. We distinguished different layers of diversity by simultaneously measuring (1) the alignment of listening patterns between individuals, and (2) within individual's breadth of music engagement.

Consistent with previous findings on the association between a population's cultural diversity and its size[10,11], we observed an increasing trend of between-level diversity with the region's population size. We add new insights using individual patterns: an individual's cultural breadth similarly expands with population size. These patterns were consistently observed across France, Brazil, and Germany. However, across these countries, we also found noticeable group-level differences for between- but not within-level diversity. This suggests that people from different countries are generally similar in their cultural breadth, but nations differ in the amount of shared repertoire among those people. We confirmed the robustness of our results through a series of control experiments that account for various biases that might stem from the data or the measure. These insights have practical implications such as guiding policy makers to distribute cultural resources across regions effectively[60].

We demonstrated how demographic factors can contribute to these observed trends. In particular, we showed a consistent cross-cultural pattern where an individual's breadth of cultural exploration tends to peak in their late 20s and narrows as they get older. This finding augments prior research that established a connection between age and diversity in cultural consumption[45,46], by providing a more detailed perspective through leveraging large-scale data that is challenging to obtain.

Using causal inference tools, we found that demographic factors largely contribute to diversity in the most populated areas. However, they did not fully account for the observed trend of increasing diversity with population size. This potentially suggests that cultural diversity found in large metropolises is not only due to demographic mixing, but also results from more frequent cultural interactions and exchanges that these larger cities enable. Interestingly, once demographic factors are accounted for, there was no discernible difference in the cultural breadth of individuals living in medium (inhabitant size around 500,000) versus large regions (more than 1.2 million), suggesting an upper limit in the extent that size fosters cultural exploration. This limit is likely reached when a region has sufficient infrastructure to support diverse cultural experiences.

Our consumer-centric perspective on cultural diversity generally contrasts with the producer-centric perspective that has been the dominant form in studies focusing on population size and culture. For instance, studies have used archaeological records and analysed how tool-making skills are copied and passed down to the next generation and, over time, accumulated as collective cultural knowledge, giving rise to more diverse cultural products[10,11]. A direct parallel comparison for such processes in the case of music would be to study whether more diverse musicl styles emerge in larger population groups, as studied by Street et al.[38]. Our study offers a novel perspective by mirroring these phenomena through the lens of consumption behaviour.

The patterns of music engagement we analysed may also reflect elements of inter-generational transmission, aligning with aspects of the ‘producer’ narrative. Studies have shown that one’s music taste is strongly influenced by their parents, translated as an increased preference for music not only from their own generation but from that of their parents[45]. In large population areas, where families exhibit more diverse music interests, this could lead to their offspring developing a more eclectic taste in music, sustaining the cultural diversity within these regions. Future studies can look to incorporate both, the production and consumption patterns, and model the interactions in the social network of individuals. This would offer a more comprehensive view of how cultural diversity is shaped, maintained, and evolves over time.

With the exception of age and gender, other demographic factors in our study were approximated from postcode level averages through inferred users’ home locations. Thus, the regional data may be limited in fully capturing the subtle variances across the individuals. Moreover, the user characteristics of the music platform may not perfectly align with the general population’s behaviour — although the general demographics appear to be similar (User demography in Methods). We also acknowledge that there may be other important confounders we did not consider and that more demographic data could ultimately change the results. Nevertheless, here we made an extensive effort to augment a variety of information that is typically overlooked, such as acquiring social connectedness and data on music venues. Our main aim in illustrating the model with a DAG is to transparently lay out our assumptions that can be refuted and corroborated by future researchers to serve as a useful framework.

Determining causal relationships would ideally necessitate intervention experiments[56,57]. However, in the context of our study and many others examining social phenomena, it is often not feasible nor ethical to design experiments manipulating individuals' cultural environments. Thus, past studies have relied on rather simple hypothetical scenarios with small sample sizes to test the causality of group size on cultural diversity and complexity[16–19,22]. Nevertheless, the advent of online experimental platforms and advanced recruitment methods now enables large-scale and cross-cultural studies. These technologies enable experiments beyond traditional paradigms by incorporating complex topological designs in the form of iterative transmissions[61–65] and studying behaviour in artificial worlds[66,67] that are closer to real-world scenarios, proving to be promising avenues. In this respect, our large observational analyses provide important insights about real-world behaviour that can go in tandem with future experimental approaches in answering questions about the general mechanisms underlying cultural evolution and transmission.

Methods

Music streams

Our research was conducted on Deezer, a globally available music streaming platform present in 187 countries, containing a catalogue of over 90 million tracks. User’s listening behaviour is captured with comprehensive logs, including the date and time of song playback, listening duration, the listener's self-reported age and gender (when provided), preferred language, type of streaming device and network connection, geographical localisation at the city-level derived from a third-party service, and whether the stream is organic (i.e., user explicitly making the choice of music) or algorithmically recommended. We gathered all songs that were streamed in France, Brazil, and Germany between 26th March and 2nd of April, 2023, a four-week period ensuring a balanced representation of weekdays and avoiding holiday seasons. To reduce noise and potential biases in our data, we excluded streams that were played for less than 30 seconds and over mobile networks, which is unreliable for geolocation tracking. The data was handed to the researchers anonymised and the analysis was not used to derive commercial profiling of any kind.

User sampling

From the sampled streams, we excluded users who streamed less than 100 times within the study window and ones who frequently travelled (more than 10 unique geo locations identified). This criterion was set by assessing the general frequency distribution of unique locations per user (Supplementary Fig. 1). The results were also replicated with more loose and strict thresholds (Supplementary Fig. 2). Each user's home location was approximated from the most frequently streamed city-level locations. Using data from GISCO (version year = 2020), French and German users were mapped to Local Administrative Unit (LAU) areas (number of LAU units: France = 34,968, Germany = 11,007). LAUs are building blocks of the NUTS (Nomenclature of territorial units for statistics) and statistical regions, and comprise the municipalities and communes of the European Statistical System (ESS). Brazilian users were mapped to municipality-level areas following the data (version year = 2021) published by the Instituto Brasileiro de Geografia e Estatística (IBGE). This mapping procedure excluded 5.4% of the users who could not be mapped to any of the regions within these geographical boundaries. Finally, we grouped the remaining users in France and Germany by the NUTS3 units (one higher-level of LAU) to reduce noise when measuring BID and WID, while keeping the municipality level for Brazil as grouping by the state was too broad. Regions with less than 200 unique users were excluded from our analysis to reduce noise and to ensure anonymity when shared as aggregated-level data. This resulted as 96 and 113 NUTS3 unit regions remaining for France and Germany respectively, and 218 municipalities of Brazil.

User demographic

In the final set, 2,544,549 users remained (France = 1,506,899 (47.1% Female), Brazil = 816,101 (33.6% Female), Germany = 221,547 (46.6% Female). A detailed demographic comparison across the countries by their self-reported age and gender and the number of monthly streaming activities can be found in Supplementary Fig. 3. To assess how representative our user sample is of the general population, we compared the user demographics to the Eurostat census data on population size, median age, and gender (Socio-demographics in Methods). We tested this for France, given the highest sample granularity and importance in our analysis involving causal testing. When compared across the regional NUTS3 population of France, strong correlations were observed between the number of users and census population size ($r = 0.90$, $p < .001$), and their median age ($r = 0.61$, $p < .001$). However, when stratified by age groups, our user sample consisted of 25% more young (between ages 15-30), 3% more mid-age (31-50), and 30% less elderly (51-80) populations. Our sample demonstrated a slightly skewed tendency towards having more male demographics (52% male) compared to France's general population (48% male). The gender ratio per region did not significantly correlate ($r = 0.13$, $p = 0.18$), suggesting platform-wise biases in gender.

Measuring diversity

Between-individual diversity (BID)

To quantify the diversity found in a given population, existing research has commonly applied measures like the Gini coefficient[68], Simpson's index[69], or Shannon's entropy[70]. However, these measures have been criticised for their arbitrary scales, making comparisons between results challenging. As a solution to this, Hill's number (also known as the effective number of species) has become a method increasingly popular to quantify the diversity of an assemblage, which allows for standardised comparisons[39,50]. In our use case, treating each song as a *specie* of a given population, we calculate between-individual diversity (BID) of music engagement as (${}^qD$), expressed as:

$$ {}^qD = \left(\sum_{i=1}^{S} p_i^q \right)^{1/(1-q)} \quad (1)$$

Here, q defines the order of Hill's number, where higher values of q emphasize the contribution of rare songs, while lower values of q focus on the abundance of popular songs. The S represents the total number of unique songs, and p signifies the relative abundance of each song.

The Hill's number of order $q = 1$ is then defined as the limit of the expression in Equation (1) as q approaches 1:

$$ {}^1D = \frac{1}{\prod_{i=1}^{R} p_i^{p_i}} = \exp\left(- \sum_{i=1}^{R} p_i \ln(p_i) \right) \quad (2)$$

which essentially becomes the exponential of the Shannon entropy in natural logarithms. In our analysis, we set the order of q to be 1, but results with other values of order q can also be found in Supplementary Table 2.

Within-individual diversity (WID)

To assess an individual's diversity of musical engagement, we employed the Generalist-Specialist Score (GS-Score), a previously validated metric in user music exploration and discovery studies[42,43]. The GS-Score of a user (u) is computed by taking the cosine similarity between a given user's mean vector $\vec{\mu}$, calculated from all the songs the user consumed within the last 28 days, and the vector representation of a randomly selected song ($\vec{s}$) they have listened to, weighted by the number of times they have listened (w_s). This vector representation is derived from the high-quality song embeddings of Deezer (Song embedding in Methods). Formally, the equation for GS-Score can be written as:

$$GS(u) = \frac{1}{\sum_s w_s} \sum_s \frac{w_s \cdot \vec{s} \cdot \vec{\mu}}{\|\vec{s}\| \cdot \|\vec{\mu}\|} \quad (3)$$

The GS-Score effectively captures the dispersion of a user's listening behaviour in the general song embedding space, whereby a 'specialist' would have a more focused interest, whereas a 'generalist' would exhibit a broader range of music engagement. To make this score consistent in the direction of our BID measure, we inverted the score (1 - GS(u)) and normalised the value to range between 0 to 100, where 100 represents maximal WID.

Embedding space

Song embedding

Many recommendation systems leverage embeddings to efficiently encode the latent relationships between users and content. This methodology ensures that closely related content is positioned near each other within the embedding space. For instance, tracks that share thematic or stylistic similarities, such as Adele's 'Someone Like You' and Sam Smith's 'Stay With Me', would be located in proximity within this space. Deezer employs the singular value decomposition (SVD)[71] technique based on the co-occurrence of millions of songs from user-generated playlists and listening patterns. This matrix is then approximated through factorisation technique to yield 128-dimensional embedding[72], capturing the nuanced relationships between songs based on user interactions and thematic links.

User embedding

Each user's listening behaviour, represented in the song embedding, can be summarised to a general position in the embedding space that defines the centre of their music engagement breadth. Similar to the song embedding, users with similar musical listening habits, often corresponding to fans of a particular genre, are positioned nearer to each other in this space. To represent each user, the mean user vector ($\vec{\mu}$) is computed by taking the average of embeddings of all songs the user listened to within the last 28 days.

Dispersion in user embedding

The frequency-based measure of between-individual diversity (Between-individual diversity (BID) in Methods) potentially overlooks the relations between the songs. An alternative approach is to measure the dispersion of regional users within the user embedding space. Regions where the users are more *spread out* are indicative of there being more diverse and unique clusters of music interests. We capture this dispersion per regional users in the embedding space as the population variance, written as:

$$\text{population variance} = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2 \quad (4)$$

where (x_i) is the cosine similarities of bootstrapped pairs of user vectors, and $(\bar{x})$ is the mean cosine similarity across all users. This is analogous to the method used for measuring WID using the GS-Score computation (Within-individual diversity (WID) in Methods).

Socio-demographics

Our DAG model includes seven socio-demographic confounders. We outline the data sources for each and discuss the rationale behind their inclusion, drawing on existing literature.

Age and gender

Studies have shown how one's musical exploration[43] and preferences[45] are influenced by age, demonstrating that one's music taste generally consolidates during adolescence. Recent studies have also shown significant differences between the consumption patterns of males and females[46,73]. We thus include the user's self-reported age and gender as potential confounders influencing music consumption diversity. Among the users who provided information (89.7%), age was computed based on the self-reported birth date. When registering, a user could specify their gender from the following options: 'Male', 'Female', 'Unknown', 'Non-Binary', 'Other', or left blank. For simplicity in our DAG analyses, we only included users with ages above 18 and below 65, and who self-identified themselves as Male or Female (13% excluded; see Supplementary Fig. 3 for demographic distributions).

Immigration, education, and income

Immigration or contact between populations has been shown to act as a vibrant channel for cross-cultural exchange, introducing fresh perspectives to a culture, and fostering innovation and complexity[10,74]. In parallel, one's educational attainment and economic status have been shown to be tightly linked with their cultural inclination. Rooted in the cultural omnivore theory widely debated in the social sciences[75], numerous studies have observed that the societal elites seek for a broader spectrum of cultural experiences[47–49]. Considering this past literature, we include dimensions that can indicate social class and immigrant status. The data were drawn from Cagé and Piketty (2023)[76], who aggregated and made openly available electoral and socio-economic data from the municipalities in France, with longitudinal records dating back as early as the 18th century (Data availability). These data are collected from the electoral reports digitised in national archives such as the L'Institut National de la Statistique et des études économiques (INSEE) — the National Institute for Statistics and Economics Studies of France. Among various socio-demographic indicators they collect, we focused on immigration, education and income at the level of communes that divide France into over 35,000 area units, which corresponds to the granularity of postcode resolution. We used these regional averages as proxies for each user's attributes (see Supplementary Fig. 9 for geographical map visualisations). Specifically, we used the most recent data from 2022 on: (1) percentage of immigrants (pimmigre2022) from the 'naticommunes' dataset, (2) percentage of residents with bachelor's degrees (pbac2022) from the 'diplomescommunes' dataset, (3) and average per capita income (revmoy2022) from the 'revcommunes' dataset (for an accurate description of the columns and source of the data, see their appendix material[76]). Assessing by the Q-Q plot, we observed that income and immigration percentages across the municipalities were highly skewed in their distribution, thus log transformation was applied (Supplementary Table 3).

Musical venues

Cultural activities concentrate in places like cities and towns[77]. Residents of metropolitan and large cities have easier access to diverse cultural offerings compared to their rural counterparts, which may subsequently influence their cultural engagement. To approximate the amount of access available to cultural events, we gathered information about the musical venues at the NUTS3 level using SongKick's database, a popular global concert discovery service. Using their API (Data availability), we initially queried 10,000 venues in France in August 2023. After excluding venues without geolocation information, 6,618 venues remained (see Supplementary Fig. 10 for geographical map visualisation).

Social connections

Engaging in international connections can significantly influence one's exposure to diverse content by providing access to cultural content beyond their own cultural sphere[78,79]. Increased international connections may also suggest extensive travel experience or a background of living abroad. Recent research has also found that individuals tend to broaden their preferences and interests towards the cultural influences of the places they visit[80]. We used the publicly available dataset released by Meta (reference period: 13th of October, 2021; Data availability) to approximate the number of international Facebook friends one has at the level of NUTS3 units. The Social Connectedness Index (SCI), first introduced by Bailey et al.[81] uses an anonymised snapshot of all active Facebook users and their friendship networks to measure the intensity of social connectedness between locations. Users are assigned to locations based on their information and activity on Facebook, including the stated city on their Facebook profile, and device and connection information. Formally, the *SCI* between two locations i and j is defined as:

$$\mathrm{SCI}_{i,j} = \frac{\mathrm{Connections}_{i,j}}{\mathrm{u}_i \times \mathrm{u}_j} \quad (5)$$

Here, u_i and u_j represent the number of Facebook users in locations i and j, and $Connections_{i,j}$ is the total number of Facebook friendship connections between individuals in the two locations. This metric effectively captures the relative probability of a Facebook friendship link between locations. To quantify the amount of international social connections, we summed the *SCI* that is paired with all other regions around the world that are not from the same country (i.e., France). We then normalised the scale and performed a log transformation to account for the skewness (Supplementary Table 1).

Causal inference

The DAG was illustrated and evaluated using the 'dagitty' R package[82]. We first checked the consistency of our data with the DAG models and the robustness of various versions of the candidate models[83,84]. Our final model passed several implied independence tests, adding credence to our DAG hypotheses. All variables were normalised with Z-scores for effect size comparisons. Variables that did not follow a normal distribution were log-transformed (Supplementary Table 1).

We used propensity scores to adjust for group differences in users living in different size regions. The propensity score condenses all observed covariates into a single metric[85]. Acting as a balancing measure, the propensity score aims to equalise the distribution of confounders between individuals across the groups. Each individual is assigned weights using inverse probability weighting (IPW)[86], which determines how much they 'contribute' to the group. Consequently, it enables the simulation of a quasi-randomised scenario to facilitate causal inference testing[56,84]. To obtain estimates of the causal effect, a weighted generalised linear model (GLM) was fitted to model the outcome of interest. To quantify the uncertainty associated with this estimate, we conducted bootstrap simulations on the entire sampling and weighting procedure (Statistical analysis in Methods).

Statistical analysis

With few expectations, all hypothesis tests were conducted using a bootstrap with 1,000 replications to derive the mean. Confidence estimates were derived from the 2.5% and 97.5% quantiles of the bootstrap means. Exceptional cases were: (1) when measuring BID and WID, where 100 unique individuals and 10,000 streams were respectively drawn without replacement. This was to balance the size of samples across the regions; (2) when causal testing, 100 simulations were run due to the heavy computational cost on computing IPW. Pearson and Spearman correlation coefficients were adjusted for multiple comparisons. One-way ANOVA was used for comparison across groups and post-hoc Tukey's test p-values were adjusted for multiple comparisons. Cohen's d was used for effect size estimates[87]. Analysis was conducted using R (version = 4.3.3) and data wrangling was done using base functions and 'tidyverse' package (version = 2.0.0).

Data availability

Aggregated BID and WID measures at the NUTS3 level are available at: https://github.com/harin-git/mus-div
Social connectedness index used for estimating international social connections is available at: https://data.humdata.org/dataset/social-connectedness-index
Eurostat census data for France and Germany are available at: https://ec.europa.eu/eurostat/web/main/data/database

Compiled longitudinal census data for France by Cagé and Piketty (2023) is available at: https://unehistoireduconflitpolitique.fr/telecharger.html
Music venues can be queried using the Songkick database through their API at: https://www.songkick.com/developer

Code availability

All analysis scripts describing the working and plottings used for the study are available at: https://github.com/harin-git/mus-div

References

1. Bettencourt, L. M. A., Lobo, J., Helbing, D., Kühnert, C. & West, G. B. Growth, innovation, scaling, and the pace of life in cities. Proc. Natl. Acad. Sci. 104, 7301–7306 (2007).
2. Bornstein, M. H. & Bornstein, H. G. The pace of life. Nature 259, 557–559 (1976).
3. Schläpfer, M. et al. The scaling of human interactions with city size. J. R. Soc. Interface 11, 20130789 (2014).
4. Luca, D., Terrero-Davila, J., Stein, J. & Lee, N. Progressive cities: Urban–rural polarisation of social values and economic development around the world. Urban Stud. 60, 2329–2350 (2023).
5. Baker, R. E. et al. Infectious disease in an era of global change. Nat. Rev. Microbiol. 20, 193–205 (2022).
6. Bettencourt, L. M. A. & Lobo, J. Urban scaling in Europe. J. R. Soc. Interface 13, 20160005 (2016).
7. Bettencourt, L. M. A. The Origins of Scaling in Cities. Science 340, 1438–1441 (2013).
8. Keuschnigg, M., Mutgan, S. & Hedström, P. Urban scaling and the regional divide. Sci. Adv. 5, eaav0042 (2019).
9. Keuschnigg, M. Scaling trajectories of cities. Proc. Natl. Acad. Sci. 116, 13759–13761 (2019).
10. Powell, A., Shennan, S. & Thomas, M. G. Late Pleistocene Demography and the Appearance of Modern Human Behavior. Science 324, 1298–1301 (2009).
11. Henrich, J. Demography and Cultural Evolution: How Adaptive Cultural Processes Can Produce Maladaptive Losses—The Tasmanian Case. Am. Antiq. 69, 197–214 (2004).
12. Kobayashi, Y. & Aoki, K. Innovativeness, population size and cumulative cultural evolution. Theor. Popul. Biol. 82, 38–47 (2012).
13. Mesoudi, A. Variable Cultural Acquisition Costs Constrain Cumulative Cultural Evolution. PLOS ONE 6, e18239 (2011).
14. Premo, L. S. & Kuhn, S. L. Modeling Effects of Local Extinctions on Culture Change and Diversity in the Paleolithic. PLOS ONE 5, e15582 (2010).
15. Shennan, S. Demography and Cultural Innovation: a Model and its Implications for the

Emergence of Modern Human Culture. Camb. Archaeol. J. 11, 5–16 (2001).
16. Derex, M., Beugin, M.-P., Godelle, B. & Raymond, M. Experimental evidence for the influence of group size on cultural complexity. Nature 503, 389–391 (2013).
17. Kempe, M. & Mesoudi, A. An experimental demonstration of the effect of group size on cultural accumulation. Evol. Hum. Behav. 35, 285–290 (2014).
18. Caldwell, C. A. & Millen, A. E. Human cumulative culture in the laboratory: Effects of (micro) population size. Learn. Behav. 38, 310–318 (2010).
19. Muthukrishna, M., Shulman, B. W., Vasilescu, V. & Henrich, J. Sociality influences cultural complexity. Proc. R. Soc. B Biol. Sci. 281, 20132511 (2014).
20. Collard, M., Buchanan, B. & O'Brien, M. J. Population Size as an Explanation for Patterns in the Paleolithic Archaeological Record: More Caution Is Needed. Curr. Anthropol. 54, S388–S396 (2013).
21. Morin, O. Reasons to be fussy about cultural evolution. Biol. Philos. 31, 447–458 (2016).
22. Fay, N., De Kleine, N., Walker, B. & Caldwell, C. A. Increasing population size can inhibit cumulative cultural evolution. Proc. Natl. Acad. Sci. 116, 6726–6731 (2019).
23. Fogarty, L., Wakano, J. Y., Feldman, M. W. & Aoki, K. The Driving Forces of Cultural Complexity. Hum. Nat. 28, 39–52 (2017).
24. Milgram, S. The Experience of Living in Cities. Science 167, 1461–1468 (1970).
25. Nilforoshan, H. et al. Human mobility networks reveal increased segregation in large cities. Nature 1–7 (2023).
26. Fearon, J. D. Ethnic and Cultural Diversity by Country*. J. Econ. Growth 8, 195–222 (2003).
27. Jayne, M. Cities and Consumption. (Routledge, London, 2005).
28. Zukin, S. Urban Lifestyles: Diversity and Standardisation in Spaces of Consumption. Urban Stud. 35, 825–839 (1998).
29. Arnould, E. J. & Thompson, C. J. Consumer Culture Theory (CCT): Twenty Years of Research. J. Consum. Res. 31, 868–882 (2005).
30. Bourdieu, P. Distinction: A Social Critique of the Judgement of Taste. (Harvard University Press, 1984).
31. Savage, P. E., Brown, S., Sakai, E. & Currie, T. E. Statistical universals reveal the structures and functions of human music. Proc. Natl. Acad. Sci. 112, 8987–8992 (2015).
32. Mehr, S. A. et al. Universality and diversity in human song. Science 366, (2019).
33. Jacoby, N. et al. Universal and Non-universal Features of Musical Pitch Perception Revealed by Singing. Curr. Biol. 29, 3229-3243.e12 (2019).
34. McDermott, J. & Hauser, M. The origins of music: Innateness, uniqueness, and evolution. Music Percept. Interdiscip. J. 23, 29–59 (2005).
35. Jacoby, N. et al. Commonality and variation in mental representations of music revealed by a cross-cultural comparison of rhythm priors in 15 countries. Nat. Hum. Behav. 1–32 (2024).
36. Rentfrow, P. J. The Role of Music in Everyday Life: Current Directions in the Social Psychology of Music. Soc. Personal. Psychol. Compass 6, 402–416 (2012).
37. Park, M., Thom, J., Mennicken, S., Cramer, H. & Macy, M. Global music streaming data

reveal diurnal and seasonal patterns of affective preference. Nat. Hum. Behav. 3, 230–236 (2019).

38. Street, S. E., Eerola, T. & Kendal, J. R. The role of population size in folk tune complexity. Humanit. Soc. Sci. Commun. 9, 152 (2022).
39. Chao, A. et al. Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. Ecol. Monogr. 84, 45–67 (2014).
40. Hsieh, T. C., Ma, K. H. & Chao, A. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). Methods Ecol. Evol. 7, 1451–1456 (2016).
41. Hill, M. O. Diversity and Evenness: A Unifying Notation and Its Consequences. Ecology 54, 427–432 (1973).
42. Anderson, A., Maystre, L., Anderson, I., Mehrotra, R. & Lalmas, M. Algorithmic Effects on the Diversity of Consumption on Spotify. in Proceedings of The Web Conference 2020 2155–2165 (ACM, Taipei Taiwan, 2020).
43. Mok, L., Way, S. F., Maystre, L. & Anderson, A. The Dynamics of Exploration on Spotify. Proc. Int. AAAI Conf. Web Soc. Media 16, 663–674 (2022).
44. Munawar, K., Kuhn, S. K. & Haque, S. Understanding the reminiscence bump: A systematic review. PLOS ONE 13, e0208595 (2018).
45. Krumhansl, C. L. & Zupnick, J. A. Cascading Reminiscence Bumps in Popular Music. Psychol. Sci. 24, 2057–2068 (2013).
46. Park, M., Weber, I., Naaman, M. & Vieweg, S. Understanding Musical Diversity via Online Social Media. Proc. Int. AAAI Conf. Web Soc. Media 9, 308–317 (2021).
47. Roose, H., van Eijck, K. & Lievens, J. Culture of distinction or culture of openness? Using a social space approach to analyze the social structuring of lifestyles. Poetics 40, 491–513 (2012).
48. Warde, A., Wright, D. & Gayo-Cal, M. Understanding Cultural Omnivorousness: Or, the Myth of the Cultural Omnivore. Cult. Sociol. 1, 143–164 (2007).
49. Flemmen, M., Jarness, V. & Rosenlund, L. Social space and cultural class divisions: the forms of capital and contemporary lifestyle differentiation. Br. J. Sociol. 69, 124–153 (2018).
50. Deffner, D., Kandler, A. & Fogarty, L. Effective population size for culturally evolving traits. PLOS Comput. Biol. 18, e1009430 (2022).
51. Premo, L. S. Effective Population Size and the Effects of Demography on Cultural Diversity and Technological Complexity. Am. Antiq. 81, 605–622 (2016).
52. Collard, M., Kemery, M. & Banks, S. Causes of Toolkit Variation Among Hunter-Gatherers: A Test of Four Competing Hypotheses. Can. J. Archaeol. J. Can. Archéologie 29, 1–19 (2005).
53. Collard, M., Buchanan, B., O'Brien, M. J. & Scholnick, J. Risk, mobility or population size? Drivers of technological richness among contact-period western North American hunter–gatherers. Philos. Trans. R. Soc. B Biol. Sci. 368, 20120412 (2013).
54. Read, D. An Interaction Model for Resource Implement Complexity Based on Risk and Number of Annual Moves. Am. Antiq. 73, 599–625 (2008).
55. Cinelli, C., Forney, A. & Pearl, J. A Crash Course in Good and Bad Controls. SSRN

Electron. J. (2020).
56. Rohrer, J. M. Thinking Clearly About Correlations and Causation: Graphical Causal Models for Observational Data. Adv. Methods Pract. Psychol. Sci. 1, 27–42 (2018).
57. Pearl, J., Glymour, M. & Jewell, N. P. Causal Inference in Statistics: A Primer. (John Wiley & Sons, 2016).
58. Greifer, N. & Stuart, E. A. Choosing the estimand when matching or weighting in observational studies. ArXiv E-Prints arXiv-2106 (2021).
59. Chatton, A. & Rohrer, J. M. The Causal Cookbook: Recipes for Propensity Scores, G-Computation, and Doubly Robust Standardization. Preprint at https://doi.org/10.31234/osf.io/k2gzp (2023).
60. García, B. Cultural policy and urban regeneration in Western European cities: lessons from experience, prospects for the future. Local Econ. 19, 312–326 (2004).
61. Harrison, P. et al. Gibbs Sampling with People. in Advances in Neural Information Processing Systems vol. 33 10659–10671 (2020).
62. van Rijn, P., Lee, H. & Jacoby, N. Bridging the prosody GAP: Genetic Algorithm with People to efficiently sample emotional prosody. in Proceedings of the Annual Meeting of the Cognitive Science Society vol. 44 (2022).
63. Anglada-Tort, M., Harrison, P. M. C., Lee, H. & Jacoby, N. Large-scale iterated singing experiments reveal oral transmission mechanisms underlying music evolution. Curr. Biol. 0, (2023).
64. Thompson, B., van Opheusden, B., Sumers, T. & Griffiths, T. L. Complex cognitive algorithms preserved by selective social learning in experimental populations. Science 376, 95–98 (2022).
65. Kirby, S., Cornish, H. & Smith, K. Cumulative cultural evolution in the laboratory: an experimental approach to the origins of structure in human language. Proc. Natl. Acad. Sci. U. S. A. 105, 10681–10686 (2008).
66. Salganik, M. J., Dodds, P. S. & Watts, D. J. Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. Science 311, 854–856 (2006).
67. Tchernichovski, O., Frey, S., Jacoby, N. & Conley, D. Incentivizing free riders improves collective intelligence in social dilemmas. Proc. Natl. Acad. Sci. 120, e2311497120 (2023).
68. Gastwirth, J. L. The Estimation of the Lorenz Curve and Gini Index. Rev. Econ. Stat. 54, 306–316 (1972).
69. Simpson, E. H. Measurement of Diversity. Nature 163, 688–688 (1949).
70. Shannon, C. E. A mathematical theory of communication. Bell Syst. Tech. J. 27, 379–423 (1948).
71. Koren, Y., Bell, R. & Volinsky, C. Matrix Factorization Techniques for Recommender Systems. Computer 42, 30–37 (2009).
72. Briand, L., Salha-Galvan, G., Bendada, W., Morlon, M. & Tran, V.-A. A Semi-Personalized System for User Cold Start Recommendation on Music Streaming Apps. in Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining 2601–2609 (Association for Computing Machinery, New York, NY, USA,

2021).
73. Christenson, P. G. & Peterson, J. B. Genre and Gender in the Structure of Music Preferences. Commun. Res. 15, 282–301 (1988).
74. Creanza, N., Kolodny, O. & Feldman, M. W. Cultural evolutionary theory: How culture evolves and why it matters. Proc. Natl. Acad. Sci. 114, 7782–7789 (2017).
75. Peterson, R. A. & Kern, R. M. Changing Highbrow Taste: From Snob to Omnivore. Am. Sociol. Rev. 61, 900–907 (1996).
76. Cagé, J. & Piketty, T. Une histoire du conflit politique. Elections et inégalités sociales en France, 1789-2022. (SEUIL, 2023).
77. Scott, A. J. A World in Emergence: Cities and Regions in the 21st Century. (Edward Elgar Publishing Ltd, Cheltenham, U.K Northampton, MA, USA, 2014).
78. Waters, J. L. Geographies of cultural capital: education, international migration and family strategies between Hong Kong and Canada. Trans. Inst. Br. Geogr. 31, 179–192 (2006).
79. Park, S. The Impact of Media Use and Cultural Exposure on the Mutual Perception of Koreans and Japanese1 The present research has been conducted by the Research Grant of Kwangwoon University in 2004. Asian J. Commun. 15, 173–187 (2005).
80. Kim, K., Askin, N. & Evans, J. A. Disrupted routines anticipate musical exploration. Proc. Natl. Acad. Sci. 121, e2306549121 (2024).
81. Bailey, M., Cao, R., Kuchler, T., Stroebel, J. & Wong, A. Social Connectedness: Measurement, Determinants, and Effects. J. Econ. Perspect. 32, 259–280 (2018).
82. Textor, J., Zander, B. van der & Ankan, A. dagitty: Graphical Analysis of Structural Causal Models. (2023).
83. Ankan, A., Wortel, I. M. N. & Textor, J. Testing Graphical Causal Models Using the R Package “dagitty”. Curr. Protoc. 1, e45 (2021).
84. Textor, J., van der Zander, B., Gilthorpe, M. S., Liśkiewicz, M. & Ellison, G. T. Robust causal inference using directed acyclic graphs: the R package ‘dagitty’. Int. J. Epidemiol. 45, 1887–1894 (2016).
85. Rosenbaum, P. R. & Rubin, D. B. The Central Role of the Propensity Score in Observational Studies for Causal Effects. Biometrika 70, 41–55 (1983).
86. Robins, J. M., Hernán, M. A. & Brumback, B. Marginal structural models and causal inference in epidemiology. Epidemiol. Camb. Mass 11, 550–560 (2000).
87. Lakens, D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front. Psychol. 4, 863 (2013).

Supplementary information

Supplementary information has been submitted along with the manuscript. It can also be viewed by following the link https://osf.io/6zugm/