nature human behaviour

LETTERS

https://doi.org/10.1038/s41562-018-0508-z

Global music streaming data reveal diurnal and seasonal patterns of affective preference

Minsu Park[1*], Jennifer Thom[2], Sarah Mennicken[2], Henriette Cramer[2] and Michael Macy[1,3*]

People manage emotions to cope with life's demands[1,2]. Previous research has identified affective patterns using self-reports[3] and text analysis[4,5], but these measures track the expression of affect, not affective preference for external stimuli such as music, which affects mood states and levels of emotional arousal[1,6,7]. We analysed a dataset of 765 million online music plays streamed by 1 million individuals in 51 countries to measure diurnal and seasonal patterns of affective preference. Findings reveal similar diurnal patterns across cultures and demographic groups. Individuals listen to more relaxing music late at night and more energetic music during normal business hours, including mid-afternoon when affective expression is lowest. However, there were differences in baselines: younger people listen to more intense music; compared with other regions, music played in Latin America is more arousing, while music in Asia is more relaxing; and compared with other chronotypes, 'night owls' (people who are habitually active or wakeful at night) listen to less-intense music. Seasonal patterns vary with distance from the equator and between Northern and Southern hemispheres and are more strongly correlated with absolute day length than with changes in day length. Taken together with previous findings on affective expression in text[4], these results suggest that musical choice both shapes and reflects mood.

Individuals manage mood to function productively and cope with the demands of daily routines[1,2]. The way in which a person chooses to regulate their mood has consequences for mental health, interpersonal functioning and personal well-being[8]; social networking, exercise and meditation generally have positive consequences, while cigarettes, drugs and alcohol can be detrimental[9]. People may also choose to regulate their mood through media consumption, including movies, TV, books and music. Among these media, music is unique in predating recorded history as a universal component of human life[10,11], one that both reflects and alters levels of emotional arousal[1,6,7], energy, wakefulness[12] and tension[1,7]. Music is also uniquely omnipresent, serving as a background soundtrack to both leisure and work activities[13], with reported listening time averaging up to 44% of waking hours[14]. While consumption of other media may also be useful for understanding emotion management, the omnipresence of music affords a singular opportunity to identify diurnal and seasonal patterns in listener's musical choices, at a very high level of temporal granularity and across diverse cultures and demographic groups.

Previous research on music consumption has relied largely on self-reports, surveys and laboratory experiments, with severely restricted numbers of participants, observation periods and geographic ranges, and without representative or naturalistic musical stimuli[14]. These limitations can now be overcome due to the rapidly growing use of mobile devices and music-streaming services worldwide. Almost half (45%) of Internet users aged 16–64 actively access licensed music throughout the day using streaming services[15] on a variety of devices, such as mobile phones, computers and smart speakers[15–17]. Of equal importance, detailed sonic and affective attributes are now available for millions of individual songs[14].

The growth of text-based social media has enabled a growing number of large-scale studies of global affect using text analysis. Recent studies used Twitter and Facebook data to take 'the pulse of the nation'[18], for cross-cultural comparisons of diurnal and seasonal patterns of positive and negative affective expressions[4], to measure affective responses to events[19] and track the consequences of shared emotionally salient news feed content[20].

Music listening differs from what people write in that it offers insight not only into what people may be feeling but also what they may want to feel. Put another way, people can choose which music to consume to achieve a desired mood (along, of course, with purposes unrelated to mood management, such as learning to sing or play the song). While previous studies of social media postings make it possible to track daily and seasonal patterns of affective expression, music consumption offers an unprecedented opportunity to identify global patterns of affective preference. Affective expression exposes others (the readers) to the writer's emotional content; conversely, the choice of music is a 'revealed preference'[21] for exposure to emotional content created by others. In short, tracking the temporal patterns of affective preference can offer a more complete picture of the emotional rhythms in human behaviour, beyond what has been learned from previous studies of affective expression.

To that end, we report hourly, daily and seasonal patterns of affective preference based on musical choices on a global scale. This descriptive account does not attempt to answer important questions about the motivations that shape listening behaviour, the emotional effects of music exposure or the latent cognitive strategies in mood management. Instead, we contribute an empirical foundation for future investigations by tracking the affective content of the music people choose to listen to, broken down by hour, day and month, and by user demographics and global locations.

Accordingly, we analysed hourly, daily and seasonal changes in affective preferences as revealed by the choice of online music streamed via Spotify around the clock across 51 countries. For each listener with at least 25 completed plays, we collected up to 1,000 completely played tracks (mean (M) = 771.9; s.d. = 336.8). The set of listeners comprised a stratified random sample of one million worldwide Spotify users, matching each country's age and gender distribution on Spotify with current data from the Central Intelligence Agency's *The World Factbook*[22]. This sample included a total of 765 million tracks played between 1 January and 31 December 2016. Completed plays measure active self-exposure

[1]Department of Information Science, Cornell University, Ithaca, NY, USA. [2]Spotify USA, Somerville, MA, USA. [3]Department of Sociology, Cornell University, Ithaca, NY, USA. *e-mail: mp867@cornell.edu; mwm14@cornell.edu

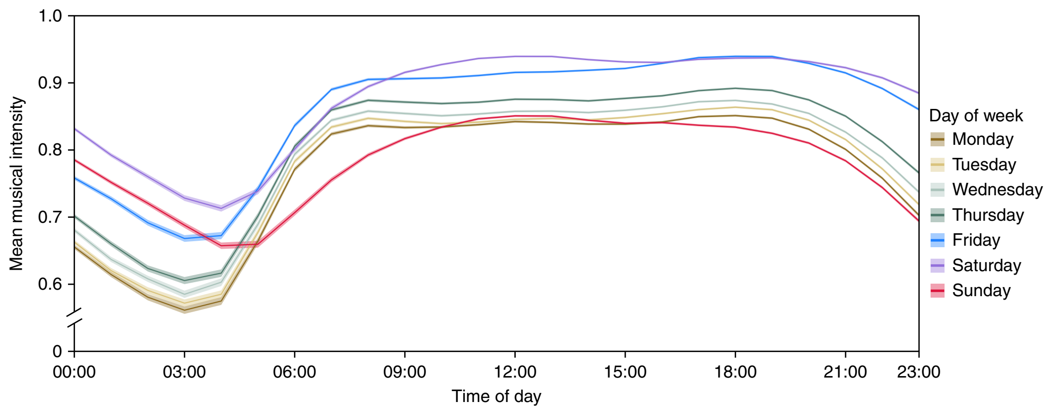

Fig. 1 | Millions of global music plays reveal diurnal affective patterns. Within-individual hourly changes in mean musical intensity scores for the global user population, broken down by day of the week, with 95% confidence intervals (translucent regions). The colours represent the days of the week, and hours were normalized to the local time (see 'Dataset description' in Methods). The *x* axis is the hour, beginning at midnight, and the *y* axis is the mean within-individual musical intensity score for each of 24 h over 7 d. The score represents the level of musical intensity among complete plays by the subset of active users during a given hour. Musical intensity levels were lowest around 03:00, with the exception of a weekend delay of 1 h (from 03:28 to 04:28), increased for about 5 h (between 03:00 and 08:00) and then were sustained for 12 h (from 08:00 to 20:00). Although the diurnal pattern was similar across all 7 d, the baseline intensity level was higher on Friday ($M = 0.879$) and Saturday ($M = 0.883$), and lower on Sunday ($M = 0.820$), compared with the other 4 d ($M = 0.828$, 0.835, 0.843 and 0.852, respectively, for Monday–Thursday; $P < 0.001$ for all pairwise comparisons). See Supplementary Table 1 for additional statistical details.

to music, excluding any songs the user may have sampled and discarded (see 'Completed plays' in the Methods for more details).

Spotify offers a way to analyse each track using 11 highly correlated audio attributes: acousticness, danceability, duration, energy, instrumentalness, liveness, loudness, mode, speechiness, tempo and valence. Principal component analysis (PCA) identified a latent construct that accounts for 29.4% of the variance in the correlation matrix (see 'Musical intensity measured by audio features of a track' in the Methods for more details). This principal component corresponds to musical intensity, ranging from highly relaxing (acoustic, instrumental, ambient, and flat or low tempo) to highly energetic (strong beat, danceable, loud and bouncy).

Aggregate temporal patterns in music consumption confound within-individual diurnal rhythms with between-individual differences in the hours when individuals with different baseline preferences for musical intensity tend to listen to music. Accordingly, we removed between-individual differences by mean-centring each individual's intensity scores such that every individual has the same baseline affective preference. We then restored between-group differences (for example, when comparing men and women or days of the week) by adding the group mean as a constant to the scores of each individual group member (see 'Measures of within- and between-individual affective preferences' in the Methods for more details). Thus, the reported temporal dynamics reflect changes over time for a prototypical group member, while differences in the intercept reflect between-group differences in baseline intensity scores.

Figure 1 reveals qualitatively identical patterns of affective preference for musical intensity on a global scale across days of the week. Musical intensity levels were highest between 08:00 and 20:00, and lowest around 03:00, with a 5-h rise (between 03:00 and 08:00) and a 7-h decline (between 20:00 and 03:00). Maximum intensity was sustained for 12 h (from 08:00 to 20:00), while minimum intensity reversed quickly and lasted only about 1 h (from 03:00 to 04:00 on weekdays and 04:00 to 05:00 on weekends). Although the timings of minimum and peak intensity were nearly identical for all 7 d, the baseline intensity level was higher on Friday and Saturday than on other days, especially in the evening when weekend social activities are likely ($M = 0.879$ and 0.883 for Friday and Saturday, compared with $0.820 < M < 0.852$ for other days; $P < 0.001$ for all pairwise comparisons; all tests for equal means throughout the paper use Welch's *t*-test to correct for unequal size and variance between paired samples; see Supplementary Table 1 for additional statistical details). The morning dip on Saturday and Sunday was delayed by 1 h (from 03:28 to 04:28), suggesting that listeners may have been sleeping in.

Overall, the diurnal pattern is remarkably similar to the temporal changes in positive affect reported in previous research using sentiment analysis of time-stamped Twitter messages[4] to measure user's affective expression. Nevertheless, we discovered one striking exception: people the world over continue to choose highly intense music throughout the day, despite the mid-afternoon slump that is registered by what they write on Twitter. The dynamic congruence with positive affect and non-congruence with negative affect suggest an intriguing hypothesis for future research: listeners select arousing music that matches their positive mood and offsets their negative mood.

Figure 2 shows that the diurnal pattern is highly consistent across five geographic regions—Latin America, North America, Europe, Oceania and Asia (Fig. 2a)—and across demographic groups based on gender (Fig. 2b), age (Fig. 2c) and chronotypes (Fig. 2d). Although the overall temporal pattern is highly robust, there are interesting between-group baseline differences. Music played in Latin America ($M = 1.053$) is relatively more intense, and music in Asia is more relaxed ($M = 0.698$) compared with Oceania ($M = 0.807$), Europe ($M = 0.804$) and North America ($M = 0.830$; $P < 0.001$ for eight pairwise comparisons of Latin America with the four other regions and Asia with the four other regions; see Supplementary Table 1 for additional statistical details). This result corroborates and extends survey- and experiment-based findings that show cultural differences in affective preferences[23]. These studies suggest that there may be cultural differences in preferences for high-arousal positive affective states, such as excitement or enthusiasm, and low-arousal positive affective states, such as calm and peacefulness, between, for example, Western and East Asian cultures.

Across the globe, intensity scores also differ by age and gender. As people get older, they listen to less-intense music ($M = 1.162$, 0.970, 0.841, 0.769 and 0.484, respectively, for the five age groups from 10–19 to over 50; $P < 0.001$ for all pairwise comparisons; see Supplementary Table 1 for additional statistical details). Intensity scores were lower for music streamed by women ($D = -0.037$;

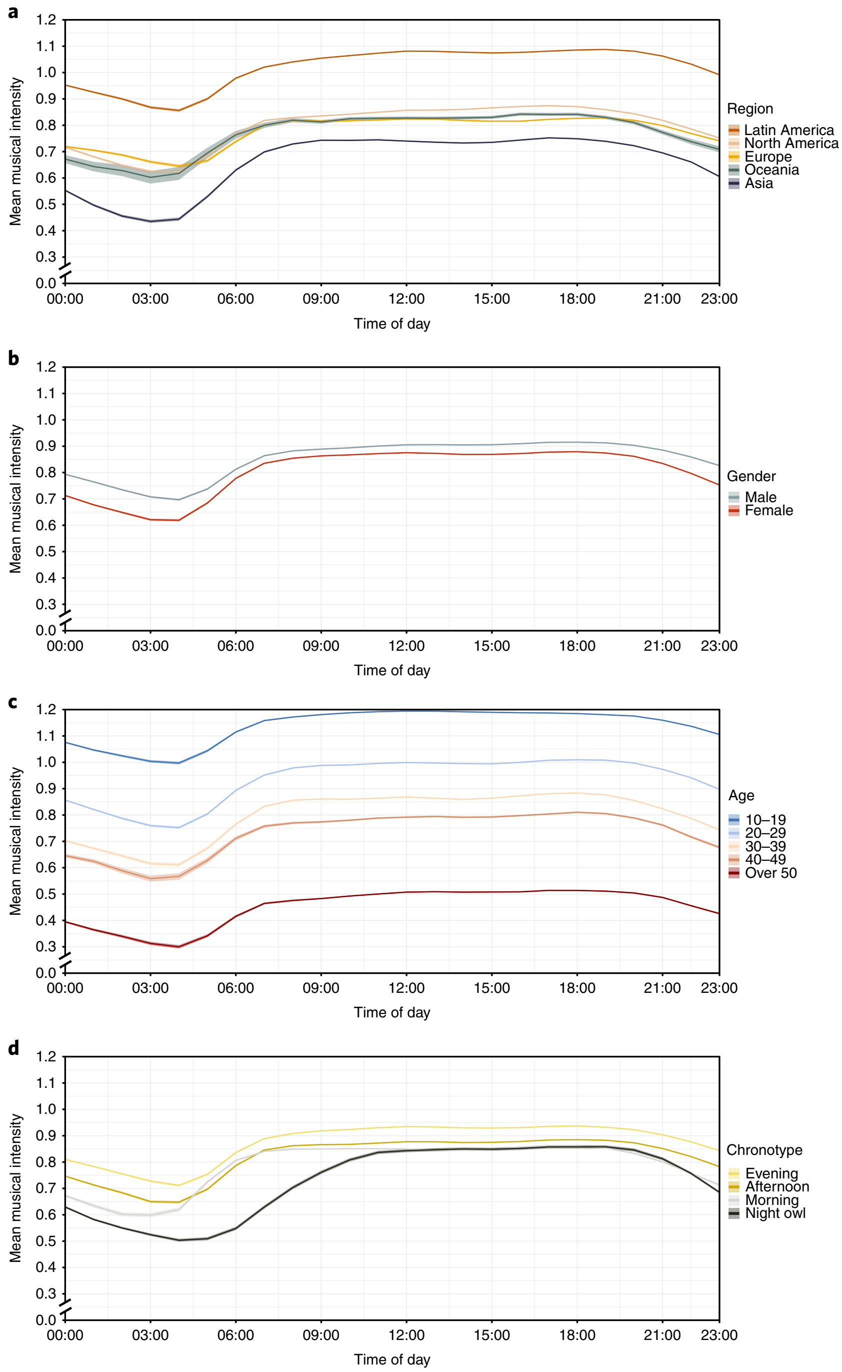

Fig. 2 | Diurnal affective patterns are robust across diverse geographic regions, demographic groups and chronotypes. a–d, Hourly within-individual changes in mean musical intensity scores across diverse geographic regions (**a**) and demographic groups based on gender (**b**), age (**c**) and chronotypes (**d**). Chronotypes are defined as 6-h intervals beginning at midnight when users are most active. The translucent regions represent 95% confidence intervals. The colours represent different groups by region, gender, age and chronotype. The *x* axis is the hour, beginning at midnight and normalized to the local time (see 'Dataset description' in Methods). The *y* axis is the mean within-individual intensity score for each of 24 h over 7 d. The score represents the level of musical intensity among complete plays by the subset of active users during a given hour. Musical intensity exhibits a dip in the morning (around 03:00), with a quick reversal and relatively constant plateau during working hours. However, there are between-group baseline differences. Across different geographical regions (**a**), music played in Latin America ($M = 1.053$) is relatively more intense and music in Asia is more relaxed ($M = 0.698$) compared with other regions ($M = 0.807$, 0.804 and 0.830 for Oceania, Europe and North America, respectively; $P < 0.001$ for all pairwise comparisons except Europe–Oceania with $P = 0.633$). In **b**, intensity scores are lower for music streamed by women ($D = -0.037$; $P < 0.001$). As people age (**c**), they listen to less-intense music ($M = 1.162$, 0.970, 0.841, 0.769 and 0.484, respectively, for the five age groups from 10–19 to over 50; $P < 0.001$ for all pairwise comparisons). Finally, night owls (**d**) listen to more-relaxing music ($M = 0.684$) than other chronotypes ($M = 0.834$, 0.861 and 0.903, respectively, for morning, afternoon and evening; $P < 0.001$ for all pairwise comparisons). Night owls also display a longer rise and larger increase in musical intensity from the morning dip to the afternoon peak ($D = 0.412$), and this temporal increase is larger among night owls compared with the increase among the other chronotypes ($D = 0.280$; $P < 0.001$). See Supplementary Table 1 for additional statistical details.

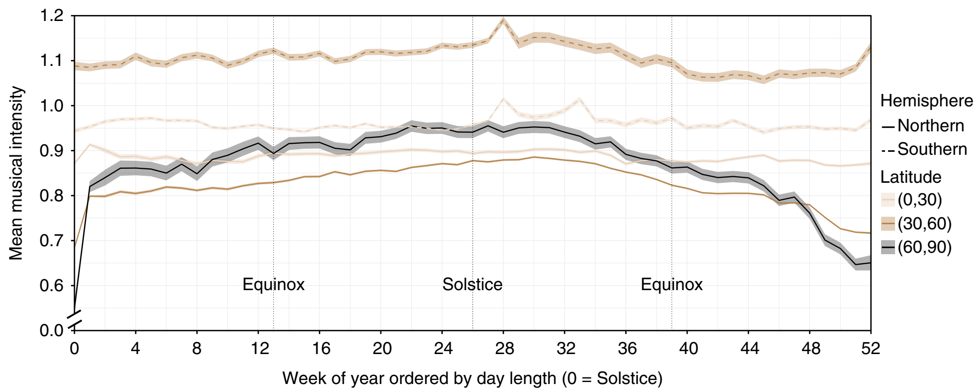

Fig. 3 | Affective preference is associated with seasonal variation in day length. Weekly changes in the mean musical intensity scores for five regions based on distance from the equator (colour) and hemisphere (line type), with 95% confidence intervals (translucent regions). Data were not available for the Southern Hemisphere at the longest distance from the equator. The *x* axis is the week of the year, ordered by day length, beginning with the winter solstice (week 0), with the summer solstice at the midpoint. Thus, the weeks on the *x* axis are different for the Southern and Northern Hemispheres (see 'Seasonal variation' in Methods). The *y* axis is the mean within-individual intensity score among complete plays by the subset of active users during each of the 53 weeks (including the 2016 leap year). Scores are broken down by distance and direction from the equator, which affect seasonal variation in day length and the timing of the winter and summer seasons. Intensity scores are highest around the summer solstice (weeks 24–28; $M = 0.919$; $P < 0.001$) and decline with day length ($r = 0.029$; $P < 0.001$), but the seasonal variation decreases with proximity to the equator. Music played around the late-December holidays is associated with a steep winter decline in intensity in the Northern Hemisphere ($D = -0.049$ for weeks 48–0 compared with other seasons; $P < 0.001$) and a sharp uptick in the Southern Hemisphere ($D = 0.087$ for week 28 compared with other seasons; $P < 0.001$). The other summer uptick in the south at latitudes under 30° S coincides with Carnival on 7 February. See Supplementary Table 1 for additional statistical details.

$t = -26.04$; *d.f.* = 1,033,792; $P < 0.001$), especially in the evening. Curiously, however, this global gender difference masks large gender differences on opposite sides of the equator, as reported in Supplementary Fig. 1a. Women stream music with higher intensity than men in the Southern Hemisphere ($D = 0.017$; $t = 6.50$; *d.f.* = 262,409; $P < 0.001$), while the pattern is the opposite in the Northern Hemisphere ($D = -0.054$; $t = -32.31$; *d.f.* = 771,029; $P < 0.001$).

The temporal dynamics are also similar across three of four chronotypes. Chronotypes were defined by when users are most actively listening, in six-hour increments beginning at midnight. The outlier group is the night owls whose baseline music intensity scores ($M = 0.684$) are lower than the scores for the other three chronotypes, with group averages increasing with the time of day during which users are most likely to listen ($M = 0.834$ for morning people, $M = 0.861$ for afternoon people and $M = 0.903$ for evening people; $P < 0.001$ for all pairwise comparisons; see Supplementary Table 1 for additional statistical details). These diurnal patterns among chronotypes closely resemble the previous findings[4] based on affect words in Twitter messages, suggesting that music consumption is closely aligned with the emotions people express. However, there is an interesting difference with affective expression in the behaviour of night owls who tend to prefer more relaxing music overall, yet display a larger increase in musical intensity during the daytime ($D = 0.412$; $t = 239.66$; *d.f.* = 2,648,000; $P < 0.001$ for the comparison between 04:00 and 18:00) compared with the daytime increase for the other 3 chronotypes ($D = 0.280$; $t = 344.11$; *d.f.* = 4,300,469; $P < 0.001$). A possible explanation is that night owls may need stronger musical stimuli to stay alert during the day.

Figure 3 reports weekly and monthly changes in music consumption that suggest that people have seasonal music preferences[24,25]. Previous research using self-reports found that listeners prefer highly arousing music during warmer months and serene music in colder seasons[25,26], but these studies were based on self-reports from small samples in specific countries. Figure 3 confirms these results on a global scale, except during winter weeks when music listening is dominated by ceremonial holiday music for Christmas and Carnival. Intensity scores peak around the summer solstice ($D = 0.078$; $t = 507.83$; *d.f.* = 107,747,995; $P < 0.001$ for the mean difference in intensity between summer weeks 24–28 and all other weeks). Intensity scores then decline with day length, but the seasonal variation decreases with proximity to the equator. Remarkably, music associated with late-December holidays is associated with a steep winter decline in intensity in the Northern Hemisphere and a sharp uptick in the Southern Hemisphere, suggesting that seasonal variation associated with holiday music can depend decisively on day length at the time of the holiday ($D = -0.049$; $t = -304.82$; *d.f.* = 116,364,849; $P < 0.001$ for winter weeks 48–0 compared with other seasons in the Northern Hemisphere; $D = 0.087$; $t = 109.51$; *d.f.* = 2,347,689; $P < 0.001$ for week 28 compared with other seasons in the Southern Hemisphere). The other summer uptick in the south at latitudes under 30° S is Carnival on 7 February.

The results in Fig. 3 resemble the seasonal patterns reported in previous studies based on affective expression in global Twitter messages[4,27]. However, while Golder and Macy[4] found that positive mood covaries with change in day length, not absolute day length, we found that absolute day length (the interval between sunrise and sunset) is a better predictor of musical intensity ($r = 0.029$; $P < 0.001$) than change in day length ($r = -0.007$; $P < 0.001$; difference in the Pearson's correlations = 0.036; Steiger's $z = 743.585$; $P < 0.001$; $n = 764{,}992{,}760$). The same result holds when excluding holiday songs ($r = 0.014$; $P < 0.001$ for absolute day length; $r = -0.008$; $P < 0.001$ for change in day length; difference in the Pearson's correlations = 0.023; Steiger's $z = 464.790$; $P < 0.001$; $n = 752{,}692{,}716$). This indicates that seasonal variations in affective music choices are more strongly influenced by seasonal activities that depend on temperature, weather, and indoor and outdoor daylight than by seasonal changes in the timing of sleep relative to the dawn signal that synchronizes the circadian pacemaker (see 'Seasonal activities and choice of music' in the Supplementary Information for additional details). Longer days are also associated with warmer temperatures, with peak temperature often lagging behind the solstice (depending on the location relative to land, water and prevailing winds). Peak music intensity also lags behind the solstice, suggesting that the mechanism that drives musical preference may be the activities associated with temperature as well as daylight.

In conclusion, data from on-demand music streaming now make it possible to study music consumption across highly diverse cultures, including countries whose music consumption is rarely studied. The findings reveal diurnal and seasonal patterns of affective preference that are highly robust across different user groups as well as countries that differ both geographically and culturally. Additional robustness tests are reported in the Supplementary Information, including seasonal patterns by different user groups (Supplementary Fig. 1), diurnal patterns broken down by day of the year (Supplementary Fig. 2), and similar results using positive and negative emotional valence instead of musical intensity (Supplementary Figs 3 and 4).

Although the robustness of the results is encouraging, there are important limitations. First, despite the reliance on a stratified random sample that reflects local census distributions of age and gender, the sample is potentially biased towards individuals who have access to streaming services and devices, particularly in lower-income countries. Second, the data are observational, and without randomized trial experiments, temporal patterns of musical intensity cannot directly test whether and when listeners use music to reflect rather than influence their mood. The relative importance of mood management and mood expression is likely to depend heavily on the cultural activities to which music provides an accompaniment, such as parties and holiday rituals.

In addition, we have data only on the intensity level of the music people choose to consume, not the affective states of the listeners. We were therefore limited to comparisons with affective expression among a different set of users on a different platform and during an earlier time period. Nevertheless, our diurnal and seasonal results show a remarkable similarity to results based on sentiment analysis of Twitter messages[4]. There are differences as well. Positive emotion in Twitter messages dips around 15:00 while the consumption of arousing music does not, suggesting that music can also be used as a mid-afternoon stimulant. While diurnal mood patterns on Twitter point to the sleep cycle as the synchronizing mechanism, listening behaviour suggests that temporal variations in preferences for affective stimuli through music may be more closely aligned with the temporal organization of daytime and night-time activities. For example, we found that listeners across the globe prefer quiet, low-intensity, relaxing music late at night and high-intensity, energetic music with a strong beat throughout the day, including late afternoon when affect expressed in writing is depressed. The comparisons suggest the possibility that music choices may reflect the emotional rhythms of daily and seasonal activities to which music contributes by shaping as well as expressing mood.

Methods

Dataset description. This study uses redacted retrospective data collected between 1 January and 31 December 2016 from music-streaming instances at Spotify—a popular streaming service for music, podcasts and video. Spotify provides 11 sonic and mood attributes (for example, acousticness, loudness, valence and energy), available through their API (https://beta.developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/). We obtained data for 764,992,760 streams from a stratified random sample of 991,035 users across 51 countries. The sample excludes users with fewer than 25 plays and was stratified to match each country's age and gender distributions and population size, based on current data from Central Intelligence Agency's *The World Factbook*[22]. The sample excludes countries where Spotify is unavailable, or with too few users after sampling to measure cross-cultural patterns. This stratified sampling adjusts the sampling frame to reflect the population distribution, since the distribution of Spotify users does not necessarily reflect the underlying population distribution. As a result, the stratified sample represents world population distribution, not Spotify user distributions over the globe. The mean age of this sample (not the service) was 37.1 years (median = 29 years; s.d. = 23.9 years) and 49.2% were female. Demographic distributions for each country can be found in Supplementary Table 2. A user's geo-location (for example, city, country, region and continent) was assigned based on the most commonly occurring geo-grid—one-tenth decimal degree by one-tenth decimal degree of pairwise latitude and longitude (approximately 100 km²)—based on Internet Protocol address. Using the Python pytzwhere library, the geo-grids were matched with time zones to normalize all time stamps to local time and adjust for daylight saving time (DST). Age and gender were obtained from current Spotify user profiles.

Chronotypes. Following Golder and Macy[4], users were allocated to four six-hour chronotypes based on the time when the user was most active on Spotify, beginning at midnight. Some 15.1% were morning people (06:00 to 12:00); 44.8% were afternoon people (12:00 to 18:00); 35.1% were evening people (18:00 to 00:00); and 5.0% were night owls (00:00 to 06:00). These chronotypes are similarly distributed across gender and age. The baseline intensity of music played by night owls differs from the other three chronotypes, as reported in Fig. 2d (see also Supplementary Fig. 5 for the distribution of plays across different times of day).

Completed plays. In contrast with radio-like streaming services, Spotify is a user-driven on-demand service with a vast catalogue from which users search for and choose songs they want to listen to. Spotify reports that more than 80% of listening on Spotify in 2016 (when we collected the data) was initiated by user selection and not through algorithmic personalization[28]. Users can also exercise selection by choosing which songs to play to completion and which to skip. We limited the analysis to completed (or non-skipped) plays to focus on the music people actively choose to listen to, excluding what they choose to skip.

Musical intensity measured by audio features of a track. Music provides listeners with an affective experience through various musical features, ranging from song lyrics to the emotional attributes of audio features. Musicologists argue that audio features (particularly biopsychological cues, such as arousal) have better cross-cultural applicability without the language constraints of lyrics[29]. Spotify's track-specific audio attribute data are considered the gold standard in music information retrieval[30]. Spotify provides 11 common audio features: acousticness, danceability, duration, energy, instrumentalness, liveness, loudness, mode, speechiness, tempo and valence (see descriptions in Supplementary Table 3). The attributes are highly correlated, and PCA identified a latent structure, with the first principal component unambiguously interpretable as a measure of intensity that explains 29.4% of the variance. We excluded the second principal component, which explained an additional 12.1% of the covariance but did not have a meaningful interpretation including shared characteristics related to known musical attribute dimensions that people usually perceive, such as arousal (similar to our intensity measure), valence and depth[31], among others. Supplementary Fig. 6 displays the locations of the 11 Spotify attributes on the PCA coordinate space for the first two principal components. Song-specific intensity scores range from −7.70 to 3.96 and are strongly associated with musical acousticness ($r = -0.765$), energy ($r = 0.867$) and valence (negative to positive emotion; $r = 0.643$; all of the Pearson's correlations are significant at $P < 0.001$; $n = 13{,}578{,}157$). Factor loadings show that tracks with high intensity tend to be fast, loud, vocal (that is, not instrumental), happy, cheerful and euphoric (see Supplementary Table 3 for the complete set of factor loadings).

Measures of within- and between-individual affective preferences. Temporal changes in affective preference were measured as the average intensity level of the music that a user listened to in each of the $24 \times 7 = 168$ h of the week. Failure to disaggregate within- and between-individual affective preferences makes changes over time uninterpretable due to the confounding of individual diurnal rhythms and temporal changes in the composition of active users on Spotify. Between-individual variation in intensity scores (that is, the average level of intensity in the music that a user listened to) captures how individuals differ from one another in their baseline affective preferences, regardless of the time of day or day of the week. Between-individual baseline intensity (BIntensity) scores were averaged over the scores for tracks played during 168 time points for each user, across all hours (which therefore does not vary from hour to hour):

$$\mathrm{BIntensity}_u = \overline{\mathrm{Intensity}_u} = \frac{1}{\|H\|} \sum\nolimits_{h \in H} \mathrm{Intensity}_u(h)$$

The within-individual intensity score (WIntensity) for a person-hour measures the signed difference between an individual's mean intensity score for that hour and their baseline score (as defined above). Within-individual differences in intensity scores measure how a given individual's affective preference varies over time, after removing differences in baseline scores between individuals who are active at different times, leaving only the change over time that is within each individual:

$$\begin{aligned}\mathrm{WIntensity}_{u,g}(h) &= \mathrm{Intensity}_u(h) - \mathrm{BIntensity}_u \\ &\quad + \frac{1}{\|UH(g)\|} \sum\nolimits_{(u,h) \in UH(g)} \mathrm{Intensity}_u(h)\end{aligned}$$

where u and h pairs indicate user-hours, and $UH(g)$ is the set of all user-hour combinations in the group g (where g can be a day of the week, region, demographic group or chronotype) for which the within-individual pattern is measured.

The final term in $\mathrm{WIntensity}_{u,g}(h)$ is the grand mean across all user-hours in g. Note that the final term is $\frac{1}{\|U(g)\|} \sum_{u \in U(g)} \mathrm{BIntensity}_u$ for groups g (such as region, demographics or chronotype) as the grand mean across all user baseline intensities in group *g*. Adding back the group-specific grand mean restores between-group differences while preserving within-individual temporal changes, since adding this constant to the mean-centred data for each individual member of that group does not affect the within-individual temporal dynamics. However, care should be taken in trying to interpret between-group differences by visual inspection of the figures, since the number of observations varies greatly over the course of the day (see Supplementary Fig. 5). Thus, a group with much higher musical intensity scores late at night (when listening is less frequent), and only slightly lower scores during the day, might have a significantly lower baseline score than might be inferred simply by imagining a horizontal line fitted to the figure.

Plots in the main text show the mean within-individual intensity scores across different groups for each of 24 h over 7 d (that is, 168 hourly observations per user):

$$\mathrm{WIntensity}_g(h) = \frac{1}{\|UG(h)\|} \sum\nolimits_{(u,g)\in UG(h)} \mathrm{WIntensity}_{u,g}(h)$$

where *u* and *g* pairs are the subset of users in group *g* who were active during hour *h*, and *UG(h)* is the set of all users in group *g* who were active during hour *h*. These scores reveal diurnal patterns in affective preferences over the course of a day.

Seasonal variation. The seasonal analysis parallels the diurnal analysis, except that intensity scores are averaged over person-weeks (or person-days for Supplementary Fig. 2) instead of person-hours. The analysis tests the hypothesized emotional effects of changing day length. The length of the day at a given location varies sinusoidally over the year, with higher amplitude waves the farther one moves from the equator, resulting in long summer days and short winter days in extreme latitudes, and consistent day length near the equator. The day length at a given location on a given day is governed by the day of the year and the latitude at that location.

Two models are widely used to estimate day length. Although the Center for Biosystems Modeling (CBM)[32] reports more accurate day length estimation than the Brock model[33] when compared with the *Astronomical Almanac*, this only applies to low and mid-latitudes, with CBM accuracy declining rapidly poleward of 60°. Therefore, we use both models, the CBM for <60° and the Brock model for ≥60°.

The Northern and Southern hemispheres have winter and summer six months apart, which makes interpretation of day length patterns awkward when the person-week (or person-day) affective preference is plotted against calendar dates. Instead, the *x*-axis in Fig. 3 is ordered by day length, starting and ending with the winter solstice, with the longest day at the mid-point. The *x* axis begins with 21 December 2016 for countries in the Northern Hemisphere and 20 June 2016 for those in the Southern Hemisphere, with the summer solstice (20 June in the north and 21 December in the south) at the mid-point, and the day preceding the winter solstice on the far right (see also Supplementary Figs 1, 2 and 4).

Group baseline comparisons. In the main text, we report baseline differences in mean musical intensity scores across groups in different group categories (for example, day of the week, age, gender, chronotype and geographical region). We performed all statistical tests of group differences in baselines using the unadjusted data, not the mean-centred data points with adjusted baselines. However, in the figures that report mean-centred within-individual results (Figs. 1–3 and Supplementary Figs 1–4), we facilitated visual inspection (both of variations around the baseline and of baseline comparisons) by adding back the mean for each group. The group means were also computed from the unadjusted data and did not reflect the mean-centring used to identify within-individual temporal variation.

Other psychological features in music attributes. Based on a hierarchical PCA on 25 computer-generated attributes for 102 music excerpts across diverse genres and styles, previous research[34] has shown that computer-generated sonic and affective features can similarly capture latent dimensions of human-perceived attributes[31] on the same music excerpts: arousal (the first principle component, indicating music that is danceable and loud), valence (the second; positive and happy) and depth (the third; instrumental and low tempo). While the arousal dimension has very similar characteristics to our intensity measure (for example, positive correlations with danceability and loudness, and negative correlations with acousticness), none of our lower-ranked PCA dimensions was directly matched with the other two dimensions. This is not surprising, given that we applied PCA to 11 audio features generated from a large body of popular music (that is, hundreds of millions of complete songs by millions of artists) while previous work relied on 25 features in hundreds of excerpts from commercially unreleased songs that were previously curated for balance across genres and styles. A curated pool of music excerpts may be suitable for the fine assessment of music preferences and validation of automated feature extraction, but the latent feature structures should not be expected to match those of actual listening behaviours.

Nevertheless, valence is included as 1 of our 11 features, and readers familiar with previous research may therefore find it interesting to see how this measure of positive and negative affect varies across time, space and demographic groups. We include results on diurnal and seasonal patterns of musical valence in the Supplementary Information (see Supplementary Results and Supplementary Figs. 3 and 4).

Effects of DST. The transition to DST provides an opportunity to tease apart the effects of day length from the potential confound of biorhythms associated with the light–dark and wake–sleep cycles. DST radically shifts the light–dark cycle, but there is only a very small change in day length, which affords the opportunity to use regression discontinuity for causal inference[35]. In our dataset, 31 countries had DST in 2016. We labelled each day of the year relative to the start and end dates for DST for a given country. For instance, Sunday 13 March 2016 was the start date of DST in the United States. Accordingly, 12 March, 13 March and 14 March were labelled −1, 0 and 1, respectively. We took mean intensity scores across 31 countries for each labelled day. For each DST start-date and end-date-based daily intensity score, we conducted two tests: (1) non-parametric discontinuity estimation using the smoothing parameter (bandwidth) proposed by Imbens and Kalyanaraman[36,37] (IK bandwidth) for discontinuity at the DST start or end dates; and (2) McCrary's test[38] for possible discontinuity around the DST start or end dates. Supplementary Fig. 7a shows the result of the non-parametric discontinuity estimation based on the start date of DST. This indicates discontinuity around New Year's Day and Christmas, but no discontinuity at the start date of DST. This was statistically confirmed by McCrary's test ($z=0.200$; $P=0.842$) and by a regression using the local approach with default IK bandwidth ($z=-1.101$; $P=0.271$; $R^2=0.144$). Supplementary Fig. 7b also shows no discontinuity at the end date of DST, which was statistically confirmed using local linear regression ($z=-0.855$; $P=0.392$; $R^2=0.399$) and McCrary's test ($z=0.195$; $P=0.846$).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Code availability
Aggregate data and code are available at https://github.com/minsu-park/affective_preference_rhythm.

Data availability
The datasets used in this study are available from Spotify, but restrictions apply to the availability of these data, which were used under an agreement for the current study, and so are not publicly available. Data are, however, available from the authors upon reasonable request and with permission from Spotify.

Received: 13 July 2018; Accepted: 3 December 2018;
Published online: 21 January 2019

References

1. Thayer, R. E., Newman, J. R. & McClain, T. M. Self-regulation of mood: strategies for changing a bad mood, raising energy, and reducing tension. *J. Pers. Soc. Psychol.* **67**, 910–925 (1994).
2. Gross, J. J. The emerging field of emotion regulation: an integrative review. *Rev. Gen. Psychol.* **2**, 271–299 (1998).
3. Clark, L. A., Watson, D. & Leeka, J. Diurnal variation in the positive affects. *Motiv. Emot.* **13**, 205–234 (1989).
4. Golder, S. A. & Macy, M. W. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science* **333**, 1878–1881 (2011).
5. Dodds, P. S. & Danforth, C. M. Measuring the happiness of large-scale written expression: songs, blogs, and presidents. *J. Happiness Stud.* **11**, 441–456 (2010).
6. Hargreaves, D. J. & North, A. C. The functions of music in everyday life: redefining the social in music psychology. *Psychol. Music* **27**, 71–83 (1999).
7. Knobloch, S. & Zillmann, D. Mood management via the digital jukebox. *J. Commun.* **52**, 351–366 (2002).
8. Gross, J. J. & John, O. P. Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *J. Pers. Soc. Psychol.* **85**, 348–362 (2003).
9. Zillmann, D. Mood management: using entertainment to full advantage. *Commun. Soc. Cogn. Affect.* **31**, 147–171 (1988).
10. DeNora, T. *Music in Everyday Life* (Cambridge Univ. Press, Cambridge, UK, 2000).
11. Sloboda, J. A. & O'Neill, S. A. in *Music and Emotion: Theory and Research* 415–429 (Oxford Univ. Press, Oxford, UK, 2001).
12. Thayer, R. E. *The Origin of Everyday Moods: Managing Energy, Tension, and Stress* (Oxford Univ. Press, Oxford, UK, 1997).
13. North, A. C., Hargreaves, D. J. & Hargreaves, J. J. Uses of music in everyday life. *Music Percept.* **22**, 41–77 (2004).
14. Greenberg, D. M. & Rentfrow, P. J. Music and big data: a new frontier. *Curr. Opin. Behav. Sci.* **18**, 50–56 (2017).

15. Ipsos Connect. *Music Consumer Insight Report* https://www.ifpi.org/downloads/Music-Consumer-Insight-Report-2016.pdf (International Federation of the Phonographic Industry, 2016).
16. *U.S. Smartphone Use* (Pew Research Center, 2016).
17. *Communications Market Report 2017: United Kingdom* (Ofcom, 2017).
18. Mislove, A., Lehmann, S., Ahn, Y.-Y., Onnela, J.-P. & Rosenquist, J. N. Pulse of the Nation: U.S. Mood Throughout the Day Inferred from Twitter *Khoury College of Computer and Information Sciences, Northeastern University* http://www.ccs.neu.edu/home/amislove/twittermood/ (2010).
19. Thelwall, M., Buckley, K. & Paltoglou, G. Sentiment in Twitter events. *J. Assoc. Inf. Sci. Technol.* **62**, 406–418 (2011).
20. Kramer, A. D. I., Guillory, J. E. & Hancock, J. T. Experimental evidence of massive-scale emotional contagion through social networks. *Proc. Natl Acad. Sci. USA* **111**, 8788–8790 (2014).
21. Samuelson, P. Consumption theory in terms of revealed preference. *Economica* **15**, 243–253 (1948).
22. *The World Factbook* (Central Intelligence Agency, 2017).
23. Tsai, J. L. Ideal affect: cultural causes and behavioral consequences. *Perspect. Psychol. Sci.* **2**, 242–259 (2007).
24. Keller, M. C. et al. A warm heart and a clear head: the contingent effects of weather on mood and cognition. *Psychol. Sci.* **16**, 724–731 (2005).
25. Pettijohn, T. F., Williams, G. M. & Carter, T. C. Music for the seasons: seasonal music preferences in college students. *Curr. Psychol.* **29**, 328–345 (2010).
26. Krause, A. E. & North, A. C. 'Tis the season: music-playlist preferences for the seasons. *Psychol. Aesthet. Creat. Arts* **12**, 89–95 (2018).
27. Harmatz, M. G. et al. Seasonal variation of depression and other moods: a longitudinal approach. *J. Biol. Rhythms* **15**, 344–350 (2000).
28. *Spotify Technology S.A. Direct Listing Prospectus* (Spotify, 2018).
29. Balkwill, L.-L. & Thompson, W. F. A cross-cultural investigation of the perception of emotion in music: psychophysical and cultural cues. *Music Percept.* **17**, 43–64 (1999).
30. Askin, N. & Mauskapf, M. What makes popular culture popular? Product features and optimal differentiation in music. *Am. Sociol. Rev.* **82**, 910–944 (2017).
31. Greenberg, D. M. et al. The song is you: preferences for musical attribute dimensions reflect personality. *Soc. Psychol. Pers. Sci.* **7**, 597–605 (2016).
32. Forsythe, W. C., Rykiel, E. J., Stahl, R. S., Wu, H. & Schoolfield, R. M. A model comparison for daylength as a function of latitude and day of year. *Ecol. Model.* **80**, 87–95 (1995).
33. Brock, T. D. Calculating solar radiation for ecological studies. *Ecol. Model.* **14**, 1–19 (1981).
34. Fricke, K. R., Greenberg, D. M., Rentfrow, P. J. & Herzberg, P. Y. Computer-based music feature analysis mirrors human perception and can be used to measure individual music preference. *J. Res. Pers.* **75**, 94–102 (2018).
35. Thistlethwaite, D. L. & Campbell, D. T. Regression-discontinuity analysis: an alternative to the ex post facto experiment. *J. Educ. Psychol.* **51**, 309–317 (1960).
36. Imbens, G. & Kalyanaraman, K. Optimal bandwidth choice for the regression discontinuity estimator. *Rev. Econ. Stud.* **79**, 933–959 (2012).
37. Gelman, A. & Imbens, G. Why high-order polynomials should not be used in regression discontinuity designs. *J. Bus. Econ. Stat.* https://doi.org/10.1080/07350015.2017.1366909 (2018).
38. McCrary, J. Manipulation of the running variable in the regression discontinuity design: a density test. *J. Econometrics* **142**, 698–714 (2008).

Acknowledgements

This research and data analysis were predominantly carried out during M.P.'s internship at Spotify. We thank M. Antoniak, F. Diaz, T. Y. Hou, J. Park, J. Zhang, members of Cornell's Social Dynamics Laboratory, and colleagues at Spotify for helpful suggestions. This research was supported by the US National Science Foundation (SES 1226483), Minerva Initiative (FA9550-15-1-0162) and the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2016S1A3A2925033). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Authors contributions

M.P. and M.M. designed the research, analysed the results, and wrote the first draft. M.P. conducted the analyses. M.P., J.T., S.M., H.C., and M.M. jointly wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41562-018-0508-z.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to M.P. or M.M.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): Minsu Park and Michael Macy

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	Spotify's internal BigQuery was used to sample and extract the initial dataset.
Data analysis	Python (2.7) and its pytzwhere (3.0) library was used to match geo-grids with time zones to normalize all time-stamps corresponding to the geo-grids to local time and adjust for daylight saving time. R (≥ 3.1) and ggplot2 (3.0.0) package were used to produce figures. R's rddtools (0.4.0) package was used for regression discontinuity tests. R's FactoMineR (1.41) was used to identify the musical intensity, the first principle component of the sonic and mood attributes of music. No other custom algorithms were developed and used and all the R scripts used to produce the results are shared in a public repository as specified in the manuscript.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

April 2018

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The datasets used in this study are available from Spotify but restrictions apply to the availability of these data, which were used under an agreement for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of Spotify.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Quantitative observational study
Research sample	The study uses 764,992,760 music streaming instances from a stratified random sample of 991,035 Spotify users across 51 countries. This stratified sampling adjusts the sampling frame to reflect world population distribution since the distribution of Spotify users does not necessarily reflect the underlying population distribution. As a result, the stratified sample represents world population distribution, not Spotify user distributions over the globe. The mean age of this sample was 37.1 (median 29; standard deviation 23.9) and 49.2% female.
Sampling strategy	The sample was stratified to match each country's age and gender distributions and population size, based on current data from CIA's World Factbook. The sample excludes countries where Spotify is unavailable or with too few users after the sampling to measure cross-cultural patterns.
Data collection	We accessed the music streaming data and audio and mood features for each individual music internally using Spotify's internal BigQuery, the Google Cloud based data warehouse. Audio and mood attributes of music are publicly available through Spotify's API (https://beta.developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/).
Timing	It is redacted retrospective data collected between January 1, 2016 and December 31, 2016.
Data exclusions	The sample excludes countries where Spotify is unavailable or with too few users (under 1,000 users) after the sampling to measure cross-cultural patterns.
Non-participation	*State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.*
Randomization	*If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.*

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging