ARTICLE

https://doi.org/10.1057/s41599-022-01139-y

OPEN

Check for updates

The role of population size in folk tune complexity

Sally E. Street ^[],^{1,2,3™}, Tuomas Eerola ^[],⁴ & Jeremy R. Kendal ^[],^{1,2,3}

Demography, particularly population size, plays a key role in cultural complexity. However, the relationship between population size and complexity appears to vary across domains: while studies of technology typically find a positive correlation, the opposite is true for language, and the role of population size in complexity in the arts remains to be established. Here, we investigate the relationship between population size and complexity in music using Irish folk session tunes as a case study. Using analyses of a large online folk tune dataset, we show that popular tunes played by larger communities of musicians have diversified into a greater number of different versions which encompass more variation in melodic complexity compared with less popular tunes. However, popular tunes also tend to be intermediate in melodic complexity and variation in complexity for popular tunes is lower than expected given the increased number of tune versions. We also find that user preferences for individual tune versions are more skewed in popular tunes. Taken together, these results suggest that while larger populations create more frequent opportunities for musical innovation, they encourage convergence upon intermediate levels of melodic complexity due to a widespread inverse U-shaped relationship between complexity and aesthetic preference. We explore the assumptions underlying our empirical analyses further using simple simulations of tune diffusion through populations of different sizes, finding that a combination of biased copying and structured populations appears most consistent with our results. Our study demonstrates a unique relationship between population size and cultural complexity in the arts, confirming that the relationship between population size and cultural complexity is domaindependent, rather than universal.

¹ Durham Cultural Evolution Research Centre, Durham University, South Road, Durham DH1 3LE, UK. ² Evolutionary Anthropology Research Group, Department of Anthropology, Durham University, South Road, Durham DH1 3LE, UK. ³ Durham Research Methods Centre, Durham University, South Road, Durham DH1 3LE, UK. ⁴ Department of Music, Durham University, Palace Green, Durham DH1 3RN, UK. ^{Sem}email: sally.e.street@durham.ac.uk

Introduction

emography, particularly population size, plays a fundamental role in cultural complexity (Mesoudi 2016). A positive correlation between population size and technological complexity is perhaps one of the most consistent and widely-discussed findings in the field of cultural evolution, supported by multiple lines of evidence (Mesoudi 2016). Theoretical models suggest that in larger populations, opportunities for innovation are more frequent and beneficial innovations are less likely to be lost due to stochastic effects, supporting diverse cultural repertoires of complex technologies (Henrich 2004; Kobayashi and Aoki 2012; Powell et al. 2009; Shennan 2001). These predictions are supported by several historical case studies, such as the substantial reduction in technological repertoire among the indigenous Tasmanian population following their isolation from mainland Australia (Henrich 2004), and the proliferation of technological innovation along with increases in population density during the Upper Palaeolithic transition in Europe and western Asia (Powell et al. 2009; Shennan 2001). Comparisons across human populations have also found positive correlations between population size and the diversity and complexity of traditional toolkits (Collard et al. 2013; Kline and Boyd 2010). Finally, transmission chain experiments have shown that larger group sizes help sustain cultural diversity and skill in artificial technological tasks (Derex et al. 2013; Muthukrishna et al. 2014).

Support for a positive correlation between population size and cultural complexity is not, however, consistent across domains. Previous work has largely focused on technological case studies, while larger populations appear to have the opposite effect on linguistic complexity. Languages with a greater number of speakers have more diverse vocabularies, but simpler grammatical structures (Bromham et al. 2015; Lupyan and Dale 2010; Nettle 2012; Reali et al. 2018), and artificial language experiments have shown that larger groups produce simpler communication systems that are easier for subsequent generations to learn and reproduce (Fay and Ellison 2013). Similarly, in the case of folk tales, population size correlates positively with diversity in tale type but negatively with diversity in narrative motifs (Acerbi et al. 2017). Taken together, these results suggest that in the linguistic domain, larger populations facilitate more frequent opportunities for innovation, but that more frequent transmission favours increased compressibility. The contrasting findings between studies of language and technology suggest that the relationship between population size and cultural complexity is not universal, but rather depends on cultural selection pressures specific to different domains (Acerbi et al. 2017; Tamariz et al. 2016).

The relationship between population size and complexity in the arts, including music, is yet to be established. In common with language and technology, larger populations may facilitate more frequent opportunities for the invention of new musical variants, increasing musical diversity. Alternatively, given that music is usually a communal activity (Savage et al. 2015; Trehub et al. 2015), larger populations may instead reduce diversity by increasing standardisation. For example, if you were to invent and sing your own version of the 'Happy Birthday' song at a social gathering, your new version would be unlikely to catch on given that the conventional melody is so well established (perhaps the best known melody worldwide; Brauneis 2009). Larger populations may, further, allow more complex melodies to emerge and spread due to greater variation in player skill level coupled with a tendency to copy prestigious, highly skilled players, in line with theoretical models based on technology (Henrich 2004). Alternatively, in common with language, more frequent transmission in larger populations may result in simplification of melodies towards more learnable forms. The cultural 'fitness landscape' for the arts is, however, distinct from those of both language and technology in that it is determined more strongly by subjective aesthetic preferences, including an enjoyment of innovation and complexity for their own sakes (Tamariz et al. 2016). Aesthetic preferences for complexity, further, appear to trade off against those for predictability and structure, given the frequently observed inverse U-shaped relationship between aesthetic appeal and complexity in the arts, including music (Chmiel and Schubert 2017; Delplanque et al. 2019; North and Hargreaves 1995; Van Geert and Wagemans 2021). Folk music scholars too have suggested that the most popular and enduring tunes are those with an optimal balance of aesthetic interest and learnability (Hillhouse 2013; Vallely 2011).

Folk 'session' tunes provide an ideal case study for investigating the relationship between population size and musical complexity. Session tunes are dance-based melodies of primarily Irish origin, such as jigs, reels and hornpipes, typically performed at informal gatherings (Gagne 1996; Vallely 2011). During a session, players take turns to 'lead' sets of tunes, with other players joining in with tunes they recognise (Gagne 1996; Vallely 2011). Importantly for the present study, session tunes are almost always performed from memory, meaning that despite increasing commercialisation and digitalisation of Irish traditional music in recent decades (Gagne 1996; Hillhouse 2013; Pendlebury 2020), individual learning and memory processes still exert considerable influence over session tune innovation and transmission. While there are tens of thousands of folk session tunes in circulation, tunes vary widely in their popularity such that only a relatively small subset enter the 'core' repertoire widely known by most session musicians (Hillhouse 2013). This means that tunes vary widely in their 'effective cultural population sizes'-i.e., the size of the population of individuals capable of performing a particular cultural variant and potentially transmitting it to others, similar to the concept of effective population size in population genetics (Kolodny et al. 2015; Powell et al. 2009; Shennan 2001). In recent decades, populations of session musicians have become globally dispersed communities comprising both local connections among musicians who play together in person as well those between geographically separated musicians linked by online media (Waldron and Veblen 2008). Session tunes also vary widely in their diversity and complexity: while some have diversified into hundreds of different versions (called 'settings'; Vallely 2011), others have only one or two known variants (Roud and Bishop 2014). Most session tunes are fairly short, highly structured and relatively easy to learn and remember, following an AABB format consisting of two repeated 8-bar sections (e.g., The Kesh, Drowsy Maggie). However, many are longer and considerably more difficult to perform, comprising 3-5 or more repeated sections and more technically challenging motifs (e.g., the Frieze Breeches, the Maid at the Spinning Wheel).

Here, we investigate the relationship between population size and complexity using analyses of a large online folk tune dataset, The Session (thesession.org; Keith 2021). The Session is a community website dedicated to Irish traditional music, used by session musicians to identify tunes, find local sessions and contribute to discussion boards. The Session collection includes a large number (>38,000) of session tunes available in ABC notation, a convenient format for extraction of measures of melodic complexity since no transcription is required. The Session is unique among online folk tune collections in that it allows users to add tunes to virtual tune collections and thus records measures of tune popularity among session musicians. We use the popularity of a tune as a proxy for its effective cultural population size —i.e., the number of players who are able to perform and potentially transmit the tune to others at real-life sessions. First,

to test the alternative predictions that larger effective population sizes increase or decrease the frequency of innovation, we investigate the relationship between tune popularity and two measures of tune diversity: richness (number of tune versions) and variation in melodic complexity across tune versions. We use several different measures of melodic complexity including tune length, entropy, and novelty. Second, to investigate whether there is a distinct relationship between population size and complexity in music compared with other domains, we investigate whether the relationship between population size and musical complexity is positive, negative or inverse U-shaped. Third, to test the prediction that larger populations facilitate stronger convergence upon preferred tune variants, we investigate the relationship between tune popularity and measures of relative diversity and evenness. Finally, to complement the empirical results, we run some simple simulations of tune diffusion through populations of varying sizes, investigating the role of social learning biases and population structure on relationships between population size and tune complexity.

Methods

Data collection. We obtained measures of popularity, diversity and complexity for a large number of folk session tunes from a weekly data dump from The Session website (https://github.com/adactio/ thesession-data). The Session functions as a platform for tune exchange and discussion among session musicians (Waldron and Veblen 2008), allowing users to both upload tunes to the collection and add tunes to their own personal tune libraries ('tunebooks'). There are often multiple versions of the same tune available on the site, known as 'settings'. Irish traditional music is a thriving and evolving genre, comprising both newly composed tunes and those that are decades or even centuries old (Hillhouse 2013). However, users of the Session site are explicitly asked to submit traditional tunes rather than their own compositions. Therefore, in theory most of the tunes on the Session website should be of traditional origin, although in practice the origin of session tunes is varied and often ambiguous, with new tunes disseminated in multiple ways, via both oral tradition and commercial distribution (Gagne 1996; Hillhouse 2013; Pendlebury 2020; Vallely 2011; Waldron and Veblen 2008). We originally obtained data at the setting level for all available tune settings (n = 38,151 settings of n = 18,042 tunes, downloaded on 01/ 02/2021).

Population size. Here, we use tune popularity as a proxy for effective cultural population size, assuming that the number of times Session website users add a tune to their virtual tunebooks is positively correlated with the size of the community of players able to perform and potentially transmit the tune at real life sessions. Popularity data are available at both the tune-level and setting-level: users can either add a tune to their tunebook (which we refer to as a 'tunebook add') or add a setting to a virtual 'set' of tunes designed to be played together back-to-back (a 'set add'). Tune-level popularity data are available only for tunes added to at least ten tunebooks (n = 9782 tunes). At the setting-level, popularity data are available for settings that have been added to at least one set (n = 13,615 settings of n = 7089 tunes, downloaded on 30/04/2021).

Tune diversity and complexity. To investigate the role of population size in tune diversity, we collapsed setting-level data to the tune-level and calculated two measures of tune diversity: tune richness, the number of unique settings per tune, and tune variation, measured as the inter-quartile range (IQR) in melodic complexity across all settings of a tune. We estimated the melodic complexity of each tune setting using seven different measures as proxies for the difficulty of learning and performing a tune from

memory: number of bars, number of notes, a measure of complexity based on melodic expectations (Eerola et al. 2006; Eerola and North 2000), a measure of complexity based on tone transitions (Simonton 1984, 1994), first order entropy of pitch class distributions, second order entropy of pitch class distributions (adjacent pitch pairs) and a measure of novelty based on melodic self-similarity (Foote 2000). The first two measures, number of bars and number of notes, capture tune length, assuming that longer tunes should generally be harder to learn and perform from memory than shorter tunes. The remaining measures capture various dimensions of melodic complexity relevant to the difficulty of learning and correctly recalling melodies. All of the melodic complexity measures have been validated in experimental studies, which find that melodies with higher scores are judged as more complex, although measures based on melodic expectations are generally most predictive of listeners' ratings (Eerola et al. 2002, 2006; Eerola and North 2000). For all melodic complexity measures, higher values indicate greater complexity and thus lower predictability and greater difficulty to learn.

To count the number of bars in each tune, we first converted tunes from ABC notation, a format commonly used in digital collections of folk tunes (Walshaw 2011), to MusicXML, which includes bar numbers, using the python script 'abc2xml' (Vree 2018). We then used string-matching functions in base R (R Core Team 2020) to extract the last bar number for each tune. Bar counts do not account for repeated sections, therefore a tune consisting of e.g., 16 bars with no repeats and a tune comprising two 8-bar repeated sections would both be counted as 16 bars long. Since session tunes typically follow an AABB structure, this approach is consistent with our study's focus on the difficulty of tunes to learn and play from memory: a tune comprising two repeated 8-bar sections (AABB format) should not be more difficult to learn and reproduce than an alternate version with two non-repeated 8-bar sections (AB format). Bar counts also do not distinguish between full and partial bars, thus anacruses (partial bars containing 'leading' notes or motifs prior to the start of the tune) are counted as an additional bar, which is reasonable given that anacruses contain unique information within the tune. We excluded a small number of settings (n = 7) which were less than four bars in length, as they were found to contain tune fragments, corrupted files or incorrectly formatted tunes (e.g., lacking bar divisions).

We obtained all other complexity measures using MIDI toolbox 1.1, a collection of MATLAB functions for extracting melodic variables from MIDI data (Eerola and Toiviainen 2016). Before reading tunes into MIDI toolbox, we converted tunes from XML to MIDI format using the Batch Convert plugin (Schmitz 2020) in MuseScore 3 (Schweer and MuseScore Developer Community 2021). The expectancy-based model (EBM) of melodic complexity attempts to capture the extent to which a melody violates regularities in Western music, taking into account tonal, intervallic and rhythmic factors (Eerola and North 2000). For example, small intervals between adjacent notes in a melody are more 'expected' than are large intervals (the principal of proximity), melodic contours are expected to be roughly symmetrical (the principal of registral return) and rhythms are expected to conform to a regular pattern of beats (Eerola and North 2000). The tone-transitions model (TTM) attempts to capture melodic originality by measuring the extent to which a melody makes use of rare intervals, based only on tone-transition frequencies (Simonton 1984, 1994). The two entropy-based measures capture the amount of information in the distribution of pitches and adjacent pairs of pitches (Eerola et al. 2006). Melodic novelty is based on self-similarity (i.e., tunes with low self-similarity are highly novel), calculated by the approach proposed by Foote (Foote 2000). In this method, the pitch-class content of a windowed segment of a tune is extracted, selfsimilarity of the pitch-class vectors is computed and the overall novelty is obtained by summing a gaussian checkerboard kernel over the diagonal of the self-similarity matrix (Eerola and Toiviainen 2016). Here, we calculate novelty at a resolution of 0.5 beats with a gaussian kernel of ten beats. Measures could not be extracted for n = 3387 settings due to corrupt or non-monophonic MIDI files. Four further settings were removed prior to analysis due to formatting issues which caused implausibly high or infinite values. After combining datasets and removing problematic settings, n = 9378 tunes and n = 12,422 settings remained with complete data on all measures for tune-level and setting-level analyses, respectively.

To investigate the extent to which larger populations facilitate convergence upon preferred tune versions, we used two different measures of convergence: variation in tune complexity relative to richness, and Pielou's evenness (Oksanen 2017). We measured variation in complexity relative to richness by fitting tune-level models in which variation (IQR) in melodic complexity across settings of a tune is predicted by both tune popularity and richness. If larger populations facilitate convergence on preferred tune variants, we expect to see that while they increase tune variation in absolute terms, they reduce variation relative to that expected based on richness. In other words, we predict that tune popularity has a negative effect on variation when controlling for richness. We calculated Pielou's evenness (Shannon's diversity divided by log richness; Oksanen 2017) based on the distribution of setting-level popularity scores ('set adds') across settings of a tune, using functions from the vegan R package (Oksanen 2017). In this context, Pielou's evenness measures the extent of skew in preferences for particular settings of a tune, where a value of 1 indicates equal preference for all tune settings, and lower numbers increasingly skewed preferences. Pielou's evenness can only be calculated for tunes with more than one setting and with popularity data available for all settings, which reduces the sample size for analyses of evenness (to n = 2380 tunes).

Data analysis. As our hypotheses concerned the effect of effective cultural population size on musical diversity and complexity, we treated diversity and complexity measures as dependent variables, predicted by popularity measures (used as proxies for population size). To investigate the effect of tune popularity on diversity, we ran models at the tune level predicting diversity measures (richness and variation in complexity) from popularity. We also included the number of days since the first version of each tune was uploaded to the Session website as an additional predictor, to control for a potential confound: over time, tunes can only increase in both diversity and popularity. In these models, we checked for collinearity between popularity and number of days uploaded by calculating variance inflation factors (VIFs). VIFs were low (<2), suggesting that the independent contribution of both popularity and time uploaded to tune diversity can be robustly estimated. Further, to ensure that this approach adequately controlled for the confounding effect of time since the tunes were uploaded to the site, we conducted simulations to investigate the likelihood of a significant positive association between tune popularity and richness under a null hypothesis of no direct causal relationship between the two, finding no increased risk of type I error (SI Section 1 and SI Figs. 1-3).

To investigate the effect of tune popularity on melodic complexity, we ran models at both the setting-level and at the tune-level. In setting-level models, we predicted tune setting complexity from the setting-level measure of popularity (number of 'set adds'), fitting tune identity as a random effect. In tune-level models, we collapsed complexity scores to the tune level by taking median values across all settings of each tune. We also explored relationships between the seven different melodic complexity measures to see if it would be appropriate to reduce dimensionality by creating a composite complexity variable. Correlations between the measures were, however, variable and generally low, and a principal components analysis did not suggest a single underlying dimension of melodic complexity (SI Section 2, SI Table 1 and SI Fig. 4). Therefore, we examined relationships between popularity and all seven measures of melodic complexity in separate analyses. Exploratory plots suggested that larger populations were associated with tunes of intermediate complexity across all measures. Therefore, prior to analysis we transformed all complexity variables to the absolute difference from the median value, such that tunes of intermediate complexity have values of zero and positive values correspond to increases or decreases in complexity relative to the median. Hence if popular tunes are intermediate in complexity, we should expect negative relationships between popularity and deviation from intermediate complexity.

Ordinary linear models were not appropriate for analysing relationships between tune popularity, diversity and complexity since the outcome variables had strongly right-skewed distributions (Crawley 2014). Exploratory analyses confirmed that the data violated key assumptions of linear regression models including linearity of relationships and normality and homoscedasticity of residuals (Hector 2015). We therefore used non-linear models to analyse the data, specifically exponential models (linear models with log-10 transformed outcome variables). We left tune evenness untransformed, however, as it had a left-skewed distribution and found instead that log-10 transforming tune popularity improved fit to linear model assumptions.

We ran all analyses in a Bayesian framework using the R package MCMCglmm (Hadfield 2010), using default prior distributions for fixed effects (normal, diffuse with a mean of zero and variance of 108) and inverse-Wishart priors (setting V = 1 and v = 0.002) for random effects and residual variance (Hadfield 2021). We ran MCMC chains for a sufficient number of iterations (25,000, with a burn-in period of 5000, sampling every ten generations) to obtain effective sample sizes of at least 500 for all parameters, and confirmed chains had converged on posterior distributions by visually examining trace plots and histograms. For all fixed effects, we report the means and 95% credible intervals from posterior distributions, along with pMCMC values (the probability that the posterior mean effect is equal to zero, such that lower pMCMC values indicate stronger evidence for effects in either direction). To measure model fit, we calculate marginal and conditional R^2 values, which indicate the proportion of variance explained by the fixed effects only, and by both the fixed and random effects together, respectively (Nakagawa and Schielzeth 2013). We also performed cross-validation to assess the models' abilities to predict data outside of the sample. Here, we repeated each model 100 times using a training dataset comprising a random sample of 75% of the dataset and assessed the model's predictive accuracy for the remaining 25% of the dataset using root mean squared error (RMSE). We then compared the distribution of RMSE values from the crossvalidation procedure with the RMSE for the original model based on the full dataset to summarise predictive performance, reporting the mean and SD for RMSE values from the crossvalidation procedure. To simplify interpretation, we normalised RMSE values by dividing them by the range of the dependent variable. Therefore, normalised RMSE values indicate the average error as a proportion of the range of the data from 0 to 1, where 0 represents perfect performance and 1 the worst possible performance. We performed cross-validation using custom functions as the MCMCglmm package does not include a builtin function for cross-validation.

Simulations. We ran some simple simulations to examine the effect of social learning bias and population structure on the relationship between population size and complexity. Here, we simulate the diffusion of tune settings in populations of musicians of varying sizes ranging from 20 to 1000 in increments of 20. Each player's initial tune setting complexity is drawn from a normal distribution (mean = 3; SD = 1), where scores of zero represent settings of intermediate complexity and negative and positive values simpler and more complex settings, respectively. The mean of the starting distribution is shifted to an arbitrary value (in this case, 3) away from zero to allow for potential movement towards settings of intermediate complexity. At each time-step, there is a small probability (0.1) that a player socially learns, i.e. copies the setting complexity of another player with a small amount of error (described below). If the player does not socially learn, there is a small probability (0.1) that they innovate by slightly altering the complexity of their own setting, otherwise, the player keeps their original setting with its complexity score unaltered. We assume low probabilities of social learning and innovation given that folk music is a fairly conservative tradition in which preservation is emphasised over individual creativity, and switching between tune versions is relatively rare.

We compare results across two different social learning conditions representing unbiased and biased social learning. Here, we aimed to find out whether a social learning bias towards intermediate levels of melodic complexity is necessary to explain our results or whether unbiased copying could also give rise to the patterns we observe in the data. Although prior evidence suggests that musicians do have a preference for tunes of intermediate complexity (Chmiel and Schubert 2017; Delplanque et al. 2019; North and Hargreaves 1995; Van Geert and Wagemans 2021), unbiased copying could potentially sustain stable traditions of tunes of intermediate complexity through regression to the mean if copying error is normally distributed. In the unbiased social learning condition, copying error is drawn from a normal distribution with a mean of zero (SD = 0.01). In contrast, in the intermediate-biased social learning condition, the mean error is shifted slightly from zero (by 0.005 or -0.005 depending on whether the individual's own previous setting is negative or positive, respectively), so that copying error is biased towards intermediate complexity scores. Here, we chose to implement learning biases through biased copying errors rather than selective preferences for copying tune settings of intermediate complexity, but we assume that both selective and transformative processes operate in real folk tune traditions and that either could potentially favour tunes of intermediate complexity. However, to explore this assumption more explicitly we run an additional set of simulations in which the preference for intermediate complexity is implemented via biased selection rather than transformation. We do so by weighting the probability that an individual is selected to copy by a negative quadratic function of their setting complexity score (i.e., weights increase as complexity scores approach zero).

We also compared effects across homogeneous and structured populations. Previous theoretical models have shown that population connectivity can be at least as important as size for maintaining complexity as both can increase effective cultural population size (Kolodny et al. 2015; Powell et al. 2009). The community of Session website users is increasingly global and diverse, with players influenced by both local musicians they encounter at real life sessions, and geographically distant musicians that they may only be exposed to through recordings or, increasingly, online (Gagne 1996; Hillhouse 2013). The relative importance of local versus global trends on individual session players' performances and preferences is, however, not yet well investigated. In the homogenous population condition, social influence is global such that players may select any individual in the population to copy. In contrast, in the structured condition, the population is subdivided into groups of ten players, and players select others to copy within their local group only. We run two further permutations of the structured population condition, one in which players are randomly assigned to groups and one in which group membership is assortative—here, we sort the population by tune setting complexity prior to division so that each group contains players with settings of similar complexities. We assume that the latter is more realistic since session players presumably wish to play with others of a similar standard to themselves.

Results

Tune popularity and diversity. We find that popular tunes are more diverse than unpopular tunes: both tune richness (number of unique settings) and variation (IQR) in melodic complexity increase with tune popularity (Table 1, SI Table 2, and Fig. 1). Normalised RMSE values for the models ranged from 0.061 to

Table 1 Results of regression model predicting tune richness (*N* settings) from popularity (*N* tunebook adds) and days since uploaded to the Session website (n = 9378 tunes, marginal $R^2 = 0.256$, normalised RMSE for full model = 0.186, mean normalised RMSE from cross-validation = 0.187, SD = 0.003).

Predictor	Post. means	l-95% Cl	u-95% Cl	рМСМС
N turacharali adda	0.000522	0.000496	0.000548	<0.001
Days since uploaded	0.000054	0.000049	0.000058	<0.001

Post. means = mean estimates from posterior distributions, I-95% CI = lower 95% credible intervals from posterior distributions, u-95% CI = upper 95% credible intervals from posterior distributions and pMCMC = pMCMC value.



Fig. 1 Scatterplot showing the relationship between richness (*N* settings) and popularity (*N* tunebook adds), across Session tunes (n = 9378). The shaded area is made up of fit lines from the entire posterior distribution, while the thick line is based on posterior mean effects. Fit lines were calculated with the time since tunes were uploaded to the site fixed at the median value.

Table 2 Results of tune-level models predicting measures of tune complexity from popularity (N tunebook adds, n = 9378).							
Outcome variable	Post. means	I-95% CI	u-95% Cl	рМСМС	Marg. R ²	RMSE	CV RMSE
Bar count	-0.000086	-0.000121	-0.000047	<0.001	0.002	0.210	0.210 [0.002]
Note count	-0.000038	-0.000077	0.000005	0.070	<0.001	0.148	0.148 [0.002]
Melodic expectations	-0.000027	-0.000033	-0.000020	<0.001	0.008	0.152	0.152 [0.003]
Tone transitions	-0.000015	-0.000024	-0.00003	0.007	0.001	0.176	0.176 [0.002]
Pitch entropy	-0.000002	-0.00003	-0.000001	<0.001	0.002	0.110	0.110 [0.002]
Pitch-pair entropy	-0.00003	-0.000004	-0.00002	<0.001	0.004	0.138	0.138 [0.002]
Novelty	-0.000005	-0.000006	-0.000004	<0.001	0.007	0.150	0.149 [0.002]

Post. means = mean estimates from posterior distributions, I-95% CI = lower 95% credible intervals from posterior distributions, u-95% CI = upper 95% credible intervals from posterior distributions, pMCMC = pMCMC value, marg. R^2 = marginal R^2 , the proportion of variance explained by the fixed effects in the models, RMSE = normalised root mean squared error for the full model and CV RMSE = mean and SD from the distribution of normalised RMSE values from the CV models. All melodic complexity measures were re-scaled as the absolute deviation from the median prior to analysis, where low values indicate tunes close to median complexity. Therefore, negative coefficients here indicate that as popularity increases, tunes move towards intermediate levels of complexity.



Fig. 2 Scatterplots showing relationships between the seven measures of melodic complexity and popularity at the tune level (*n* = 9378). Measures of complexity have been re-scaled so that they represent absolute deviation from the median values, i.e., lower values indicate tunes closer to intermediate complexity while higher values tunes of higher or lower complexity. Shaded areas represent fit lines from the entire posterior distribution, while thick lines are based on posterior means.

0.228 (corresponding to mean prediction errors of ~6–23% of the range of the data) and were almost identical to those calculated from the cross-validation procedure, suggesting that models had good out-of-sample predictive ability.

Tune popularity and complexity. In tune-level analyses of popularity and melodic complexity, we find negative effects of tune popularity on deviation from median complexity, i.e., tunes of intermediate complexity are most popular, consistently across all complexity measures (Table 2 and Fig. 2). The strongest effects are found for measures of complexity based on melodic expectations and melodic novelty, while number of notes has the weakest association with popularity (Table 2). Marginal R^2 values are, however, very low (<0.1), suggesting that many other factors besides popularity affect melodic complexity (Table 2). Normalised RMSE values ranged from 0.110 to 0.210 and were almost identical to mean RMSE values calculated from the cross-validation procedure, suggesting good out-of-sample predictive ability (Table 2).

In setting-level analyses of popularity and melodic complexity (with tune identity as a random effect), we find similar although less consistent results compared with tune-level analyses. Popular settings have intermediate levels of melodic complexity in terms of bar count, melodic expectations and pitch-pair entropy (Table 3 and SI Fig. 5). Marginal R^2 values were very low (≤ 0.001), while conditional R^2 values far higher (0.20–0.70), indicating strong tune-level random effects on setting complexity (Table 3). Normalised RMSE values were lower for setting-level compared with tune-level analyses of popularity and complexity (0.056–0.119), suggesting better within sample predictive ability. However, mean RMSE values from the cross-validation procedure were slightly higher than those from the full setting-level models, suggesting that the setting-level models had lower out-of-sample predictive ability compared with the tune-level models (Table 3).

Tune popularity and convergence. When controlling for tune richness, tune popularity is negatively related to variation in melodic complexity (SI Table 3), suggesting that although larger effective cultural populations increase tune diversity in absolute terms, they also facilitate greater convergence upon preferred levels of tune complexity and thus reduce relative diversity. This interpretation is supported by a negative effect of popularity on Pielou's evenness (Table 4 and Fig. 3).

Table 3 Results of setting-level models predicting measures of tune complexity from popularity (N set adds, n = 12,422).

Outcome variable	Post. means	I-95% CI	u-95% Cl	рМСМС	Marg. R ²	Cond. R ²	RMSE	CV RMSE
Bar count	-0.000530	-0.000898	-0.000179	0.003	<0.001	0.515	0.117	0.181 [0.002]
Note count	0.000037	-0.000428	0.000495	0.860	<0.001	0.434	0.097	0.139 [0.002]
Melodic	-0.000146	-0.000193	-0.000094	<0.001	0.001	0.701	0.056	0.111 [0.003]
expectations								
Tone transitions	-0.000075	-0.000190	0.000029	0.187	<0.001	0.495	0.102	0.155 [0.002]
Pitch entropy	-0.000001	-0.000009	0.000009	0.913	<0.001	0.617	0.054	0.094 [0.002]
Pitch-pair entropy	-0.000012	-0.000021	-0.000001	0.020	<0.001	0.551	0.067	0.108 [0.002]
Novelty	-0.000011	-0.000026	0.000005	0.170	<0.001	0.202	0.119	0.141 [0.002]

Post. means = mean estimates from posterior distributions, I-95% CI = lower 95% credible intervals from posterior distributions, u-95% CI = upper 95% credible intervals from posterior distributions, pMCMC = pMCMC value, marg. R^2 = marginal R^2 , the proportion of variance explained by the fixed effects in the models, cond. R^2 = conditional R^2 , the proportion of variance explained by the fixed affects. RMSE = mean and SD from the distribution of normalised RMSE values from the CV models. All melodic complexity measures were re-scaled as the absolute deviation from the median prior to analysis, where low values indicate tunes close to median complexity. Therefore, negative coefficients here indicate that as popularity increases, tunes move towards intermediate levels of complexity.

Table 4 Results of regression model predicting Pielou's evenness from log10-transformed popularity (*N* tunebook adds) and days since uploaded to the site (n = 2380 tunes, marginal $R^2 = 0.386$, normalised RMSE for full model = 0.171, mean normalised RMSE from cross-validation = 0.171, SD = 0.004).

Parameter	Post. means	I-95% CI	u-95% Cl	рМСМС			
Log10 N	-0.282296	-0.300604	-0.267657	<0.001			
Days since uploaded	0.000021	0.000015	0.000028	<0.001			
Post. means = mean estimates from posterior distributions, I-95% CI = lower 95% credible intervals from posterior distributions, u-95% CI = upper 95% credible intervals from posterior distributions and pMCMC = pMCMC value.							

Consistency of effects across time periods. Exploratory plots showed that tunes uploaded early in the Session website's history (~2001-2005) are extremely popular, far more so than would be expected if tunes gradually accumulate in popularity over time (SI Fig. 1). This pattern may be driven partly by biases in the curation of the Session tune collection: early users may have prioritised uploading highly popular tunes from the 'core' session repertoire, with more obscure tunes taking longer to be included. Therefore, we explore consistency of effects across different time periods in the Session website history by re-running our analyses on two subsets of the data split at the year 2005. We find that tune diversity increases with popularity across both time periods (SI Tables 4 and 5). Relationships between popularity and intermediate complexity at the tune level are limited to older tunes, apart from effects of popularity on melodic novelty which are retained in both time periods (SI Table 6), while at the setting level, effects of popularity on complexity are more consistent between the time periods (SI Table 7). In terms of convergence, we find that tune popularity reduces relative variation in tune complexity across both time periods for entropy and noveltybased measures, but only among older tunes for other measures (SI Table 8), while popularity is associated with a reduction in evenness in both time periods (SI Table 9).

Simulation. Across all simulation conditions, tune richness (number of unique settings) increases linearly with population size, due to the increased frequency of innovation in larger populations (Fig. 4). Mean complexity remains constant across population sizes, centred either on the initial population mean (3) if copying is unbiased, or shifted towards 0 if copying is biased towards intermediate complexity (Fig. 4). Larger populations



Fig. 3 Scatterplot showing the relationship between Pielou's evenness and popularity (*N* tunebook adds), across Session tunes (n = 9378). The shaded area is made up of fit lines from the entire posterior distribution, while the thick line is based on posterior mean effects. Fit lines were calculated with the time since tunes were uploaded to the site fixed at the median value.

have greater variation (SD) in complexity in absolute terms (Fig. 4), which is expected due to sampling effects. However, relative to richness, variation in complexity declines with population size, particularly when populations are subdivided into local groups (regardless of whether groups are randomly or assortatively assigned, Fig. 4). Therefore, in line with our empirical findings, our simulations suggest that larger populations support more innovation and diversification of tunes, but reduce relative variation in complexity, particularly when populations are structured. Unbiased social learning can sustain intermediate levels of complexity in tune traditions, but intermediate-biased social learning is required for initially simple or complex tunes to gravitate towards intermediate complexity over time. Figure 5 illustrates typical trends in complexity over time for small (n = 100) versus large (n = 1000) populations across the simulation conditions. When populations are homogenous, variation in complexity declines over time as populations converge upon either the starting mean (under unbiased copying) or 0 (under intermediate-biased copying). In contrast, structured populations maintain variation over time, unless overridden by the effect of



Fig. 4 Scatterplots illustrating effects of population size on tune diversity and complexity across simulation conditions. Scatterplots illustrating the relationship between population size and richness (1000 s of unique settings), mean complexity, SD complexity and SD complexity relative to (divided by) richness, across the six simulation conditions: global populations with unbiased copying, global populations with intermediate-biased copying, randomly selected local populations with intermediate-biased copying, assortatively selected local populations with intermediate-biased copying, assortatively selected local populations with intermediate-biased copying.



Fig. 5 Plots illustrating representative trends in individual simulations for small (n = 100) versus large (n = 1000) populations. Plots illustrating typical trends in tune complexity (where 0 = intermediate and positive and negative scores higher and lower complexity, respectively) over time in **a** small (top row, *n* = 100) and **b** large (bottom row, *n* = 1000) simulated populations across the six simulation conditions: global populations with unbiased copying, global populations with intermediate-biased copying, randomly selected local populations with unbiased copying, randomly selected local populations with unbiased copying and assortatively selected local populations with intermediate-biased copying.

biased copying. In all conditions, larger populations sustain more variation in complexity (in absolute terms) over time compared with smaller population, but the effect of population size on relative variation depends on population structure. In homogenous populations, smaller populations lose more variation in complexity over time than larger populations, likely due to stochastic effects. However when populations are structured and copying is biased, larger populations lose more variation over time since they have more to begin with. When implementing the preference for intermediate complexity through selective copying rather than biased copying error, we find that while selection results in faster convergence towards intermediate complexity than transformation, the results are otherwise qualitatively comparable (SI Figs. 6 and 7).

Discussion

We find that popular folk session tunes are more diverse, but converge more strongly upon intermediate levels of melodic complexity, compared with obscure tunes. The positive correlation we identify between tune popularity and diversity is consistent with the hypothesis that tunes played by larger communities of musicians have more frequent opportunities to diversify due to innovation or copying errors. Therefore, larger effective cultural population sizes support more diverse cultural repertoires in folk music as they do in other domains including technology and language (Acerbi et al. 2017; Bromham et al. 2015; Collard et al. 2013; Henrich 2004; Kline and Boyd 2010; Kobayashi and Aoki 2012; Lupyan and Dale 2010; Nettle 2012; Powell et al. 2009; Reali et al. 2018; Shennan 2001). Together with our simulations which consistently showed that larger populations increase tune diversity as a result of more frequent innovation, these results suggest that the positive relationship between population size and innovation rate generalises across diverse contexts. The inverse-U shaped relationships that we generally identify between popularity and melodic complexity, however, differ from both the positive and negative relationships previously identified in studies of technology and language respectively (Acerbi et al. 2017; Bromham et al. 2015; Collard et al. 2013; Kline and Boyd 2010; Lupyan and Dale 2010; Nettle 2012; Reali et al. 2018). The reductions in relative diversity and evenness that we find in popular tunes suggest that larger populations facilitate stronger convergence upon favoured tune variants, in this case those with intermediate levels of complexity. While effects of tune popularity on diversity were consistent across older and more recent tunes, some effects of popularity on complexity were limited to older tunes, which tend to be more popular. Tunes may therefore need sufficient time (15+ years in the present dataset) to gain enough popularity to result in strong convergence towards intermediate levels of melodic complexity. Taken together, our results identify a distinct effect of population size on complexity in the musical domain, confirming that the relationship between population size and complexity is not universal, but rather depends on selective pressures specific to different cultural domains (Acerbi et al. 2017; Tamariz et al. 2016).

Our findings suggest that a detailed understanding of how cultural fitness landscapes vary between different domains, particularly in terms of the costs and benefits of innovation, is essential for understanding the cultural evolution of complexity. Efficiency generally trumps aesthetic concerns in the domains of language and technology, but innovation is less constrained in the arts. In support of this explanation, a transmission chain study showed that innovation and increases in complexity were most frequent when an experimental task was framed as an aesthetic rather than technological or linguistic challenge (Tamariz et al. 2016). The optimal balance between familiarity and innovation in artistic domains will depend, further, on distinct features of specific artistic genres and social contexts. While improvisation is important in, for example, jazz and North Indian classical music, Irish traditional music is generally more conservative, with individual innovation limited by an emphasis on preservation (Vallely 2011) and by the constraints of typical folk tune structure, phrasing and tonality (Savage et al. 2020). The finding that session tunes of intermediate complexity are most popular is consistent with previous evidence that preferences for novelty in music trade off against those for predictability (Chmiel and Schubert 2017; Delplanque et al. 2019; North and Hargreaves 1995; Van Geert and Wagemans 2021), confirming the suggestion of folk scholars that tunes with an optimal balance of playability and aesthetic interest are most likely to become established favourites among session players (Hillhouse 2013; Vallely 2011). While our simulations showed that intermediate levels of tune complexity can be maintained by unbiased learning over time, a bias towards intermediate complexity is required for initially complex or simple tunes to consistently shift towards intermediate levels of complexity. In our simulations, we assumed that biases towards intermediate folk tune complexity operate through social learning, but acknowledge that other factors could potentially favour intermediate levels of tune complexity, including biases in innovation and social coordination.

Our results suggest that larger cultural population sizes do not necessarily increase or decrease complexity but rather increase diversity, providing greater potential for selection towards optimal levels of complexity for a given trait. In this way, our empirical results are consistent with classic population genetics models, which show that larger populations generate more variation and thus potential for stronger selection towards the optimum phenotype for a given environment, while stochastic processes dominate in smaller populations (Bromham et al. 2015; Collard et al. 2005, 2013; Shennan 2001). Our simulations, however, suggest that biological evolutionary mechanisms do not necessarily generalise to all cultural evolutionary contexts. When we assume that social influence is global, small populations actually lose more variation over time compared with larger populations (Fig. 5), likely because in smaller populations there is less innovation to compensate for global convergence caused by social learning. However, the reverse is true when we model structured populations. Local influence sustains more variation in populations than global influence as it prevents variation collapsing into a small number of dominant 'traditions', and since larger populations have more variation to begin with, they have more to lose, at least when copying is biased towards intermediate levels of complexity. Although the relative importance of local versus global trends on session musicians is not well understood, individuals are presumably more strongly influenced by the musicians they encounter most frequently at real-life sessions and therefore the local social influence condition likely best reflects real-life population dynamics among session musicians. Therefore, taken together, the empirical and simulation results are consistent with more recent genetic models suggesting that the strength of selection does not always increase with effective population size, but rather depends on various contextual factors including underlying mutation processes and population dynamics (Lanfear et al. 2014).

Evolutionary approaches have a long and controversial history in the study of folk music. While early folk song researchers, such as Cecil Sharp and Béla Bartók, viewed folk song traditions in explicitly evolutionary terms (Bennett 2016; Pendlebury 2020), ethnomusicologists have now largely abandoned evolutionary approaches in response to twentieth century abuses of evolutionary theory to justify racist and colonialist beliefs about the cultural superiority of Western music (Savage 2019). However,

researchers in the modern field of cultural evolution completely reject Spencerian, progressivist models of evolution which have no basis in biological reality (Mesoudi 2011), and a renewed interest in cultural evolutionary approaches to music has recently begun to emerge (e.g., Mehr et al. 2019; Ravignani et al. 2016; Savage 2019). Even outside the field of cultural evolution, Irish traditional music is often discussed in implicitly evolutionary terms. For example, the Companion to Irish Traditional Music (Vallely 2011) describes how new tunes enter the core repertoire as follows: "Communities of traditional musicians tend to vote collectively with their fingers. In a largely unspoken process of selection, a minority of tunes possessed of that special combination of playability and aesthetic interest gradually unfold themselves into the traditional repertoire[...]". Our findings fit exactly with this characterisation and show that a preference for melodies of intermediate complexity, previously demonstrated in laboratory-based experimental studies (Chmiel and Schubert 2017; Delplanque et al. 2019; North and Hargreaves 1995; Van Geert and Wagemans 2021), affects large-scale naturally-occurring trends in folk music. We therefore demonstrate that a cultural evolutionary framework can make a useful contribution to the study of folk music and hope that our findings will generate more interest in cultural evolutionary approaches among folk scholars. Further investigations of the role of population size in complexity across multiple, cross-cultural folk music corpora would be of great value in understanding the extent to which psychological factors shaping trends in folk music vary across cultures (Jacoby and McDermott 2017).

Although we identify patterns consistent with causal effects of population size on complexity in music in our empirical data, we must acknowledge that our statistical analyses are correlational in nature and therefore that alternative explanations are possible. In particular, popularity may be influenced by melodic complexity rather than the other way around. Given the commonly inverse-U-shaped relationship between melodic complexity and musical appeal (Chmiel and Schubert 2017; Delplanque et al. 2019; North and Hargreaves 1995; Van Geert and Wagemans 2021), users of the Session website likely prefer to add tunes of intermediate complexity to their tunebooks. Therefore, our results are consistent both with a causal effect of effective cultural population size on folk tune complexity and of folk tune complexity on effective cultural population size. In fact, changes in complexity and effective population size could reinforce one another in a runaway process, a possibility that has not yet been investigated in theoretical analyses (Kolodny et al. 2015). Similarly, while we focus on the possibility that that larger communities of players create more frequent opportunities for tunes to diversify, the reverse causal scenario is possible due to sampling biases (Acerbi et al. 2017)-popular tunes may have more versions recorded on the Session than unpopular tunes because more effort is made to catalogue their diversity. Tune age may also have effects on popularity that we are unable to account for in our analyses, as detailed historical data on the origin of the vast majority of folk session tunes are not available. Finally, there is some uncertainty about the extent to which the tune popularity measures we obtain from the Session are good proxies for the effective cultural population size of each tune. Presumably, when a user adds a tune to their virtual tune book or a setting to a virtual set, they like the tune and have some interest in performing it, but it does not necessarily mean that they will actually learn and potentially transmit it to others. Nonetheless, it is encouraging to observe the same general relationships between tune diversity, complexity and popularity as the empirical results in the simulations in which we explicitly model effects of population size, rather than popularity. We hope that our findings will inspire future experimental studies on the role of population size in musical

complexity, allowing for more in-depth investigation of causal processes.

We have shown that popular folk tunes are more diverse than less popular tunes, and typically intermediate in complexity. These results suggest that tunes played by larger communities of musicians, i.e., those with larger effective cultural population sizes, have more frequent opportunities to diversify and converge more strongly towards an optimum balance of playability and aesthetic interest. The relationship between population size and cultural complexity for music is distinct from that found previously for both language and technology, most likely due to the relative lack of functional constraints on innovation within the arts. The present results conform to suggestions that fundamental cultural evolutionary processes operate differently across domains (Nettle 2017; Sperber 1996; Tamariz et al. 2016). Therefore, an understanding of how cultural evolutionary dynamics vary across domains is crucial for a more generally applicable theory of cultural evolution and researchers should not assume that processes identified in studies of technology or language will necessarily generalise to other domains, particularly the arts. This study, together with a welcome recent proliferation of cultural evolutionary studies in music (e.g., Mehr et al. 2019; Ravignani et al. 2016; Savage 2019), highlights the importance of considering the full spectrum of human cultural activities for a comprehensive understanding of cultural evolution.

Data availability

Data and code associated with this study are available in the DataDryad repository at the following link: https://doi.org/10. 5061/dryad.rv15dv48h.

Received: 14 October 2021; Accepted: 21 March 2022; Published online: 28 April 2022

References

- Acerbi A, Kendal J, Tehrani JJ (2017) Cultural complexity and demography: the case of folktales. Evol Hum Behav 38(4):474–480. https://doi.org/10.1016/j. evolhumbehav.2017.03.005
- Bennett J (2016) Béla bartók's evolutionary model of folk music. Twent Century Music 13(2):291–320. https://doi.org/10.1017/S1478572216000062
- Brauneis R (2009) Copyright and the world's most popular song. J Copyr Soc USA 56(2-3):335-426. https://doi.org/10.2139/ssrn.1111624
- Bromham L, Hua X, Fitzpatrick TG, Greenhill SJ (2015) Rate of language evolution is affected by population size. Proc Natl Acad Sci USA 112(7):2097–2102. https://doi.org/10.1073/pnas.1419704112
- Chmiel A, Schubert E (2017) Back to the inverted-U for music preference: a review of the literature. Psychol Music 45(6):886–909. https://doi.org/10.1177/0305735617697507
- Collard M, Kemery M, Banks S (2005) Causes of toolkit variation among huntergatherers: a test of four competing hypotheses. Can J Archaeol 29(1):1–19. http://www.jstor.org/stable/41103514
- Collard M, Ruttle A, Buchanan B, O'Brien MJ (2013) Population size and cultural evolution in nonindustrial food-producing societies. PLoS ONE 8(9):e72628. https://doi.org/10.1371/journal.pone.0072628
- Crawley, M. J. (2014). Statistics: an introduction using R, 2nd edn. Wiley, New York.
- Delplanque J, De Loof E, Janssens C, Verguts T (2019) The sound of beauty: How complexity determines aesthetic preference. Acta Psychol 192:146–152. https://doi.org/10.1016/j.actpsy.2018.11.011
- Derex M, Beugin M-P, Godelle B, Raymond M (2013) Experimental evidence for the influence of group size on cultural complexity. Nature 503(7476):389–391. https://doi.org/10.1038/nature12774
- Eerola T, Himberg T, Toiviainen P, Louhivuori J (2006) Perceived complexity of western and African folk melodies by western and African listeners. Psychol Music 34(3):337–371. https://doi.org/10.1177/0305735606064842
- Eerola T & North, AC (2000) Expectancy-based model of melodic complexity. Proceedings of the 6th International Conference on Music Perception and Cognition. Department of Psychology, Keele University. pp.1177-1183

Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1. 110.3186

- Eerola T, Toiviainen P. (2016) MIDI toolbox 1.1. Retrieved from https://github. com/miditoolbox/1.1
- Eerola T, Toiviainen P, Krumhansl CL (2002) Realtime prediction of melodies: continuous predictability judgements and dynamic models. Proceedings of the 7th international conference on music perception and cognition. Adelaide, SA. pp. 473-476. Retrieved from http://www.cc.jyu.fi/~ptee/ publications/2_2002.pdf
- Fay N, Ellison TM (2013) The cultural evolution of human communication systems in different sized populations: usability trumps learnability. PLoS ONE 8(8):e71781
- Foote J (2000) Automatic audio segmentation using a measure of audio novelty. Proceedings of IEEE international conference on multimedia and expo. IEEE, New York. pp. 452–455. Retrieved from https://ieeexplore.ieee.org/document/ 869637
- Gagne R (1996) From session to CD: recent influences on Irish traditional dance music in America. Folklore Forum 27(2):79-95. Retrieved from https://hdl. handle.net/2022/2224
- Hadfield JD (2010) MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. J Stat Softw 33(2):1–22. https://doi.org/ 10.18637/jss.v033.i02
- Hadfield JD (2021) MCMCglmm course notes. Retrieved from https://cran.rproject.org/web/packages/MCMCglmm/vignettes/CourseNotes.pdf
- Hector A (2015) The new statistics in R: an introduction for biologists. Oxford University Press, Oxford
- Henrich J (2004) Demography and cultural evolution: how adaptive cultural processes can produce maladaptive losses—the Tasmanian case. Am Antiquity 69(2):197–214. https://doi.org/10.2307/4128416
- Hillhouse A (2013) Hooks and new tunes: contemporary Irish dance music in its transnational context. Ethnomusicol Ireland 2/3:38–60. Retrieved from http:// www.ictm.ie/wp-content/uploads/2013/08/ICTM-journal-Hillhouse.pdf
- Jacoby N, McDermott JH (2017) Integer ratio priors on musical rhythm revealed cross-culturally by iterated reproduction. Curr Biol 27(3):359–370. https:// doi.org/10.1016/j.cub.2016.12.031
- Keith J (2021) The session. https://thesession.org. Accessed 21 Jan 2021.
- Kline MA, Boyd R (2010) Population size predicts technological complexity in Oceania. Proc R Soc B 277(1693):2559–2564. https://doi.org/10.1098/rspb. 2010.0452
- Kobayashi Y, Aoki K (2012) Innovativeness, population size and cumulative cultural evolution. Theor Popul Biol 82(1):38–47. https://doi.org/10.1016/j.tpb. 2012.04.001
- Kolodny O, Creanza N, Feldman MW (2015) Evolution in leaps: the punctuated accumulation and loss of cultural innovations. Proc Natl Acad Sci USA 112(49):E6762–E6769. https://doi.org/10.1073/pnas.1520492112
- Lanfear R, Kokko H, Eyre-Walker A (2014) Population size and the rate of evolution. Trends Ecol Evol 29(1):33–41. https://doi.org/10.1016/j.tree.2013.09.009
- Lupyan G, Dale R (2010) Language structure is partly determined by social structure. PLoS ONE. https://doi.org/10.1371/journal.pone.0008559
- Mehr SA, Singh M, Knox D, Ketter DM, Pickens-Jones D, Atwood S, Lucas C, Jacoby N, Egner AA, Hopkins EJ, Howard RM, Hartshorne JK, Jennings MV, Simson J, Pinker S, O'Donnell TJ, Krasnow MM, Glowacki L (2019) Universality and diversity in human song. Science 366(6468):eaax0868. https:// doi.org/10.1126/science.aax0868
- Mesoudi A (2011) Cultural evolution. Chicago University Press, Chicago
- Mesoudi A (2016) Cultural evolution: a review of theory, findings and controversies. Evol Biol 43(4):481-497. https://doi.org/10.1007/s11692-015-9320-0
- Muthukrishna M, Shulman BW, Vasilescu V, Henrich J (2014) Sociality influences cultural complexity, Proc R Soc B 281:20132511. https://doi.org/10.1098/rspb. 2013.2511
- Nakagawa S, Schielzeth H (2013) A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods Ecol Evol 4(2):133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x
- Nettle D (2012) Social scale and structural complexity in human languages. Philos Trans R Soc B 367(1597):1829–1836. https://doi.org/10.1098/rstb.2011.0216
- Nettle D (2017) What is cultural evolution like? Retrieved from https://www. danielnettle.org.uk/2017/10/30/hotte-8-cultural-evolution/
- North AC, Hargreaves DJ (1995) Subjective complexity, familiarity, and liking for popular music. Psychomusicology 14(1-2):77-93. https://doi.org/10.1037/ h0094090
- Oksanen J (2017) Vegan: ecological diversity. R Package Version 2.4-4, 11. https:// cran.r-project.org/package=vegan
- Pendlebury C (2020) Tune families and tune histories: melodic resemblances in British and Irish folk tunes. Folk Music J 11(5):9684.
- Powell A, Shennan S, Thomas MG (2009) Late Pleistocene demography and the appearance of modern human behavior. Science 324(5932):1298–1301. https://doi.org/10.1126/science.1170165

- R Core Team (2020) R: a language and environment for statistical computing. Version 3.6.3. R Foundation for Statistical Computing, Vienna http://www.rproject.org/
- Ravignani A, Delgado T, Kirby S (2016) Musical evolution in the lab exhibits rhythmic universals. Nat Hum Behav 1(0007):1–7. https://doi.org/10.1038/ s41562-016-0007
- Reali F, Chater N, Christiansen MH (2018) Simpler grammar, larger vocabulary: How population size affects language. Proc R Soc B 285:20172586. https://doi. org/10.1098/rspb.2017.2586
- Roud S, Bishop J (2014) The New Penguin Book of English Folk Songs. Penguin Classics, London
- Savage PE (2019) Cultural evolution of music. Palgrave Commun 5(1):16. https:// doi.org/10.1057/s41599-019-0221-1
- Savage PE, Brown S, Sakai E, Currie TE (2015) Statistical universals reveal the structures and functions of human music. Proc Natl Acad Sci USA 112(29):8987–8992. https://doi.org/10.1073/pnas.1414495112
- Savage PE, Chiba G, Currie TE, Suzuki H, & Atkinson Q (2020) Sequence alignment of folk song melodies reveals cross-cultural mechanisms of musical evolution. PsyArXiv https://doi.org/10.31234/osf.io/5rj6y
- Schmitz J (2020) Batch convert Musescore Plugin. Retrieved from https:// musescore.org/en/project/batch-convert
- Schweer W, MuseScore developer community (2021) MuseScore 3.6.2. https://musescore.org/en
- Shennan S (2001) Demography and cultural innovation: a model and its implications for the emergence of modern human culture. Camb Archaeol J 11(1):5–16. https://doi.org/10.1017/s0959774301000014
- Simonton DK (1984) Melodic structure and note transition probabilities: a content analysis of 15,618 classical themes. Psychol Music 12(1):3–16. https://doi.org/ 10.1177/0305735684121001
- Simonton DK (1994) Computer content analysis of melodic structure: classical composers and their compositions. Psychol Music 22(1):31–43. https://doi. org/10.1177/0305735694221003
- Sperber D (1996) Explaining culture: a naturalistic approach. Blackwell Publishing, Oxford
- Tamariz M, Kirby S, Carr JW (2016) Cultural evolution across domains: language, technology and art. Proceedings of the 38th annual conference of the cognitive science society, Cognitive Science Society, Philadelphia, United States. pp. 2765-2770. Retrieved from https://researchportal.hw.ac.uk/en/ publications/cultural-evolution-across-domains-language-technology-and-art
- Trehub SE, Becker J, Morley I (2015) Cross-cultural perspectives on music and musicality. Philos Trans R Soc B 370(1664):20140096. https://doi.org/10. 1098/rstb.2014.0096
- Vallely F (2011) The companion to Irish traditional music. Cork University Press, Cork
- Van Geert E, Wagemans J (2021) Order, complexity, and aesthetic appreciation. Psychol Aesthet Creat Arts 14(2):135–154. https://doi.org/10.1037/ aca0000224
- Vree W (2018) abc2xml. Retrieved from https://wim.vree.org/svgParse/abc2xml. html.
- Waldron JL, Veblen KK (2008) The Medium is the Message: Cyberspace, Community, and Music Learning in the Irish Traditional Music Virtual Community. Journal of Music, Technology & Education 1(13):99–111. https://doi. org/10.1386/jmte.1.2and3.99_1
- Walshaw, C. (2011). The abc music standard 2.1. Retrieved from http://abcnotation.com/wiki/abc:standard:v2.1.

Acknowledgements

We are grateful to Jeremy Keith, founder of thesession.org, for assistance accessing the Session folk tunes database.

Author contributions

SES conceived the study, designed the research, collected the data, analysed the data and led on manuscript writing. TE advised and assisted with statistical analysis and contributed to manuscript writing. JRK advised on statistical analysis, led the simulation analyses and contributed to manuscript writing. All authors have approved the submitted version and agree to be accountable for their contributions.

Ethical approval

Not applicable—this article does not contain any studies with human participants performed by any of the authors.

Informed consent

Not applicable—this article does not contain any studies with human participants performed by any of the authors.

ARTICLE

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1057/s41599-022-01139-y.

Correspondence and requests for materials should be addressed to Sally E. Street.

Reprints and permission information is available at http://www.nature.com/reprints

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/ licenses/by/4.0/.

© The Author(s) 2022