

Neurocase

The Neural Basis of Cognition

ISSN: 1355-4794 (Print) 1465-3656 (Online) Journal homepage: <http://www.tandfonline.com/loi/nncs20>

Measuring the representational space of music with fMRI: a case study with Sting

Daniel J. Levitin & Scott T. Grafton

To cite this article: Daniel J. Levitin & Scott T. Grafton (2016) Measuring the representational space of music with fMRI: a case study with Sting, *Neurocase*, 22:6, 548-557, DOI: [10.1080/13554794.2016.1216572](https://doi.org/10.1080/13554794.2016.1216572)

To link to this article: <http://dx.doi.org/10.1080/13554794.2016.1216572>



View supplementary material [↗](#)



Published online: 12 Aug 2016.



Submit your article to this journal [↗](#)



Article views: 2773



View related articles [↗](#)



View Crossmark data [↗](#)

Measuring the representational space of music with fMRI: a case study with Sting

Daniel J. Levitin ^a and Scott T. Grafton ^b

^aDepartment of Psychology, McGill University, Montreal, Canada; ^bDepartment of Psychological and Brain Sciences, University of California at Santa Barbara, Santa Barbara, CA, USA

ABSTRACT

Functional brain imaging has revealed much about the neuroanatomical substrates of higher cognition, including music, language, learning, and memory. The technique lends itself to studying of groups of individuals. In contrast, the nature of expert performance is typically studied through the examination of exceptional individuals using behavioral case studies and retrospective biography. Here, we combined fMRI and the study of an individual who is a world-class expert musician and composer in order to better understand the neural underpinnings of his music perception and cognition, in particular, his mental representations for music. We used state of the art multivoxel pattern analysis (MVPA) and representational dissimilarity analysis (RDA) in a fixed set of brain regions to test three exploratory hypotheses with the musician Sting: (1) Composing would recruit neural structures that are both unique and distinguishable from other creative acts, such as composing prose or visual art; (2) listening and imagining music would recruit similar neural regions, indicating that musical memory shares anatomical substrates with music listening; (3) the MVPA and RDA results would help us to map the representational space for music, revealing which musical pieces and genres are perceived to be similar in the musician's mental models for music. Our hypotheses were confirmed. The act of composing, and even of imagining elements of the composed piece separately, such as melody and rhythm, activated a similar cluster of brain regions, and were distinct from prose and visual art. Listened and imagined music showed high similarity, and in addition, notable similarity/dissimilarity patterns emerged among the various pieces used as stimuli: Muzak and Top 100/Pop songs were far from all other musical styles in Mahalanobis distance (Euclidean representational space), whereas jazz, R&B, tango and rock were comparatively close. Closer inspection revealed principled explanations for the similarity clusters found, based on key, tempo, motif, and orchestration.

ARTICLE HISTORY

Received 18 April 2016
Accepted 19 July 2016

KEYWORDS

Music cognition;
neuroimaging; case studies;
MVPA; mental imagery

Our understanding of the structure and function of the human brain has been informed over the past two decades by unprecedented advances in neuroimaging technology, particularly by fMRI applied in controlled experiments. Prior to this, nearly all of our understanding of functional neuroanatomy came from case studies, and these remain a prime source of information about the links between brain and behavior. Although case studies lack experimental control – no two lesions are perfectly identical – dozens of case studies have become famous for what they've taught us, such as those of Phineas Gage, Solomon Shereshevsky (Luria's "S"), Louis Victor Leborgne (Broca's Tan), and Henry Molaison (Scoville & Milner's HM). Case studies are not deliberate the way that controlled experiments are. They are opportunistic in that they tend to be the result of either workplace accidents, nature's cruelty, or the surgeon's scalpel, chosen by history not because they were the ideal studies but because they came to the attention of a medical doctor or research scientist and provided clues to neural function that could not be obtained in any other way.

One of the biggest unsolved puzzles in contemporary cognitive neuroscience concerns the nature of expertise, and this has rarely been investigated in case studies even though experts are


almost by definition unique outliers (notable exceptions are Luria's mnemonist, "S", and the autopsy of Albert Einstein's brain). With the creation of a massive literature characterizing brain function by fMRI, it is becoming increasingly feasible to attempt case studies of remarkable individuals using functional as well as anatomic imaging. Clearly, there are some humans who are remarkably better at things than everyone else, and their expertise shows up in a variety of domains such as chess, athletics, science, literature, politics, art, and music. How do experts attain their skill and creativity? To what extent is their achievement based on differences in neuroanatomy or genetic propensities? For expert musicians there are relevant questions concerning the mental representations of music, and how perception, cognition, and memory interact.

The study of music within the cognitive neurosciences offers a compelling avenue for exploring these questions. Music is one of a small set of human universals (Brown, 1991) and there exists wide variation in musical ability within the population. Musical behaviors activate regions throughout the brain, bilaterally, including cortical and subcortical regions, the pons, brain stem, and cerebellum (Levitin & Tirovolas, 2009). Musical expertise requires an exquisite orchestration

CONTACT Daniel J. Levitin  daniel.levitin@mcgill.ca

Present address of Daniel J. Levitin is Haas School of Business, University of California, Berkeley, CA, USA.

This article was originally published with errors. This version has been corrected. Please see Corrigendum (<http://dx.doi.org/10.1080/13554794.2016.1272293>).

 The supplemental data for this article can be accessed [here](#)

© 2016 Informa UK Limited, trading as Taylor & Francis Group

of brain networks and systems, including those for sensation, perception, decision-making, motor control, memory, attention, problem solving, emotion, and categorization (Levitin, 2007, 2012).

At the heart of musical accomplishment is the ability to manipulate in mind rich representations of the desired soundscape. Melodies and rhythms can be considered as auditory objects, possessing certain properties such as transformation invariance (Levitin, 1999; White, 1960). At the dawn of modern psychology, Noah Porter (1871) defined mental representation as “the power to recall, represent, and to reknow objects which have been previously ... presented to the mind. Thus, I gaze upon a tree, a house, or a mountain... I close my eyes, and ‘my mind makes pictures when my eyes are shut.’ I at once represent or see with ‘my mind’s eye’ that which I saw just before with the eyes of the body... In sense-perception and consciousness, the mind *presents* to itself for the first time the objects of its direct and original knowledge. In *representation*, it presents these objects a second time, or re-presents them” (c.f. Attneave & Curlee, 1977; Shepard, 1984; Tolman, 1926).

Mental representations are not always veridical or accurate depictions of the physical world, as noted for the domains of time (Maskelyne and Wundt, cited in Boring, 1923; Jaśkowski, 1999; Treisman, 1963), color (De Valois & De Valois, 1997; Shepard & Cooper, 1992), and spatial orientation (Tversky, 1981). Their imperfections can help inform us about what is salient in the environment, and can inform evolutionary questions about the most ecologically relevant features of everyday stimuli.

In contrast to these other stimuli, representations of music retrieved from memory have been shown to preserve many of the details of heard music. Participants who produced songs from memory (by singing, humming or whistling) tended to do so at the correct pitch (Levitin, 1994) and tempo (Levitin & Cook, 1996), a feat not reducible to motor memory. Brief snippets of music can be recognized by timbre alone, even when played backwards, or at a duration so short, 100–400 ms, that no rhythm or melody has time to develop (Krumhansl, 2010; Schellenberg, Iverson, & Mckinnon, 1999). Participants unselected for musical ability were able to accurately identify familiar melodies from pitch or rhythm alone (Houlihan & Levitin, 2011), suggesting that a single feature or dimension can activate a mental representation that is satisfied with a “nearest exemplar” kind of heuristic (e.g., Hongwen, Hebert, Efros, & Kanade, 2015; Zhang, Furtlehner, & Sebag, 2008). Our mental representations for music have been shown to incorporate certain regularities of the physical world such as self-similarity, as evidenced by the ability to model the rhythmic structure of compositions spanning 400 years, across styles and genres, with a $1/f$ distribution (Levitin, Chordia, & Menon, 2012).

It is popular lore that Mozart’s composing style resembled dictation, that he composed as quickly as he could write down the parts (Vigeland, 2009; Zaentz & Forman, 1984). The strong version of this view has been discredited (Keefe, 2006; Konrad, 1992; Solomon, 1995). Much of Mozart’s compositional work was done mentally before writing anything down, but what he wrote down tended to be a short-hand or sketch of the work. Melody and harmony notes would be filled in at a later time, possibly according to rules or his own stylistic conventions,

such that the process of editing and writing merged with the process of composing.

Other composers have reported that the process of writing music is difficult to understand or explain, but that the music arrives in many cases fully- or mostly fully-formed (Levitin, 2009). This raises the question of the relationship between imagination and perception. Several studies have compared listening and imagining music in participants unselected for musical ability (Bunzeck, Wuestenberg, Lutz, Heinze, & Jancke, 2005; Halpern & Zatorre, 1999; Yoo, Lee, & Choi, 2001; Zatorre & Halpern, 2005; Zatorre, Halpern, Perry, Meyer, & Evans, 1996). The principal finding is that auditory imagery appears to draw on most of the neural structures used in auditory perception. However, in contrast to visual imagery, in which the early visual cortex seems to be activated, there is no evidence that area A1 is activated during auditory imagery.

Kraemer, Macrae, Green, and Kelley (2005) used fMRI to study unprompted imagery during silent gaps they imposed in popular music. They found a functional dissociation between imagining familiar and unfamiliar music in the left primary auditory cortex (A1), and that imagining familiar music produced greater activation in auditory association areas. In contrast, *listening* to familiar music produced lesser activation in auditory association areas, and activity in left A1 did not differ as a function of familiarity. Kleber, Birbaumer, Veit, Trevorrow, and Lotze (2007) studied professional opera singers with fMRI during overt singing and imagined singing and found that overt singing activated predominantly areas processing complex motor sequences and sensory feedback/control, and that most of these areas related to motor processing were also active during imagery. Consistent with previous research, authors found no involvement of primary auditory cortex during imagery whereas secondary auditory areas showed activation. Imagined singing activated a large frontoparietal network, indicating increased involvement of working memory processes during mental imagery.

The present study applies relatively new methods for mapping different aspects of the representational space for music including composition, retrieval and imagery of known songs and perception of both familiar and unfamiliar music. To accomplish this requires involvement of an expert with an uncommon clustering of skills. Ideally, one would want to find a musician who is at once an expert composer and instrumentalist, and who is equally comfortable across a broad range of musical styles. Here, we used fMRI to acquire functional brain scans of the musician Sting. Sting is a multi-instrumentalist (bass, guitar, lute, vocals) and composer (melodist and lyricist) whose work spans many different genres and styles, including pop, rock, classical, and jazz. He has won 16 Grammy Awards. Given the importance of genre in characterizing musical preferences (Rentfrow, Goldberg, & Levitin, 2011; Rentfrow et al., 2012), we were curious to know how different genres of music would be represented in the brain of an expert whose work spans many genres. One might imagine that lyric music would be treated separately from instrumental music, that music from historical eras would be treated separately from contemporary music (due to changes in harmonic and rhythmic norms).

Given the prior work on auditory imagery and listening, summarized above, we further sought to explore the extent to which perception and imagery are similar in a musician who is, according to his own accounts, able to imagine the music he composes quite vividly.

Rather than performing direct comparisons between scans acquired during different aspects of music perception and composition, we used state of the art multivoxel pattern analysis (MVPA) and representational dissimilarity analysis (RDA) in a fixed set of brain regions to test three exploratory hypotheses: (1) whether the components during the composition of music would induce similar multivoxel patterns relative to non-musical creative processes, (2) whether the representations for a song that was perceived would be different from the same song represented by imagery, and (3) whether the cortical representations of music from different genres would cluster in terms of genre, suggesting a similarity metric for how the musician/composer's brain organizes them.

Methods

Participant

Participant was a 55-year old right-handed male, with normal hearing and no history of neurological disorders. He was recruited through personal association with one of the authors, and served without pay. He provided informed consent both to the experiment and to the release of the data herein. A film crew documented the experiment (Pochmursky, 2009) and released it as a television documentary on PBS Nova (in the US), CTV (in Canada), and the National Geographic (in Europe, Latin America, and Asia) networks. Preliminary informal analyses using the general linear model and *post hoc* contrasts were performed at that time, which included identification of networks possibly involved in composing, listening, and imagining music.

Stimuli

There were no external stimuli for Experiment 1, during which the participant was instructed to compose in his head. For Experiments 2 and 3, we carefully chose selections that we knew S would be familiar with, so as not to confound familiarity with novelty. In addition, we selected three soundalike pop hits that were novel to distinguish familiarity of a genre with familiarity of a particular composition.

Full resolution .wav soundfiles were extracted from commercial compact discs (CDs), or downloaded from the Getty Images website (see supplementary material for full citations of the stimuli used). We then volume normalized the files, using the procedure recommended in Caclin, McAdams, Smith, and Winsberg (2005). Five participants adjusted the volumes of the experimental stimuli, so that they sounded subjectively equivalent to them, while listening through headphones with a similar transfer function to the MR compatible headphones in the fMRI facility, and at a listening level equivalent to that we used in the scanner, 75 dB(A). We took the median of these ratings and applied gain reduction or amplification as required, using the *gain*

function in ProTools (Avid, Daly City, CA). This method is regarded as more accurate than RMS or other electronic methods for normalization, because the latter can be biased by brief peaks or dips in the soundfile's amplitude profile. The resulting loudness normalized files were re-encoded as .wav files and transferred to CD for presentation at the experimental session.

The participant listened through MR-compatible headphones with an attached MR-compatible microphone (Commander XG headset, Resonance Technology, Inc., Northridge, Calif., USA).

Imaging methods

After informed consent was obtained, the scanning was conducted in a single session at the Montreal Neurological Institute's (MNI) Brain Imaging Unit (BIC). The scans were conducted on a 3 Tesla Siemens Trio with a 12-channel head coil. Functional scans were acquired with a sparse sampling echo-planar imaging (EPI) sequence that eliminated acoustical noise artifacts during auditory presentations (repetition time: 30 s, echo time: 30 ms; flip angle: 90°, matrix size 64 x 64, 2D axial slice series, field of view of 25.6 cm, slice thickness of 4 mm). Thus, each experiment consisted of a set of 30 s TRs where there was a 27.5 s period of constant gradient application (during which the stimulus or instruction to imagine or compose were presented) followed by a 2.5 s period with gradient application and EPI sampling. After completion of the functional scans a T1 weighted magnetization prepared rapid gradient echo (MPRAGE) anatomic scan was acquired with a time repetition/time echo of 2100 ms/2.4 ms, flip angle of 8 degrees, time to inversion of 1100 ms, matrix size of 256 x 256, 3D sagittal slice series, field of view of 22 cm, slice thickness of 1 mm. The duration of the MPRAGE scan was 9' 42".

Procedure

Upon arrival at the BIC, the participant (S) was briefed on the procedures, filled out a standard questionnaire asking about metal in and on the body, and assured he could discontinue or pause the experiment at any time. For each the following conditions detailed below, S had 27.5 s of scanner silence to complete the requested tasks, followed by a scan pulse sequence of 2.5 s. S was given a squeeze ball to indicate distress or desire to pause or make a comment. Experimenter checked with S regularly to ensure his comfort and willingness to continue.

Experiment 1: composing

Compose global condition. S was instructed to "compose a fragment of a song you've not thought of before. Ideally, it would have lyrics – just a line or two of melody, rhythm and lyrics, the first thing that pops in your head." S reported during the experiment that he did not opt to generate lyrics.

Compose melody condition. S was instructed to think only about the melody.

Compose rhythm condition. S was instructed to think only about the rhythm of the song.

Compose melody & rhythm together condition. S was instructed to think only about the melody and rhythm.

Silent baseline condition. S was instructed to “clear your mind and try not to think about anything in particular.” (S is an advanced practitioner of meditation and yoga and found this easy to do.)

Prose control condition. S was instructed to write a piece of prose, as though writing a speech or prepared remarks.

Visual art (canvas) control condition. S was instructed to imagine painting a canvas, to picture the different colors and shapes and where he would place things on the canvas.

Replications. The above instructions were given again, in random order, for three replication blocks. Replication 1: visual art, compose global, compose rhythm, compose rhythm, compose lyrics, compose melody, rest, compose prose. Replication 2: prose, compose global, compose lyrics, compose melody, compose rhythm, visual art, rest. Replication 3: compose global, compose melody, compose lyrics, compose rhythm, compose prose, rest, visual art.

Experiment 2: listening versus imagining

The Experimenter played a soundfile not contained in the stimulus set (“My Favorite Things” by Rogers & Hammerstein, performed by the Blue Monkey Project) so that S could request adjustments to make the music play back at a comfortable listening level.

Participant was instructed as follows: “When I say the name of the song, imagine that it is playing in your head as vividly as possible, from the beginning of the song. If you don’t know the song, just tell me and we’ll move on to a different one. Then, after you’ve imagined it for 25 seconds, you’ll hear the scanner make its eh-eh-eh noise, and then I’ll actually play you the music. Just listen carefully to it as it plays.” S imagined, and then listened to, the selections listed in Table 1, in the order listed. After the first four selections, we administered a silent 30 s baseline and then resumed the playlist.

Experiment 3: listening to different genres

S was instructed simply to listen to music for Experiment 3, and that it would come from different genres of music. S listened to the selections listed in Table 1, in the order listed.

Image analysis

Preprocessing

DICOM images were converted to NIFTI format and then preprocessed using SPM12 (<http://www.fil.ion.ucl.ac.uk/spm/>) operating on the MATLAB software (Mathworks, Waltham, MA). Images within each run were realigned to each other and the three runs were aligned using rigid body transformation. All aligned scans were resliced to native space using

Table 1. List of all stimuli used in the experiment.

Condition/Artist	Genre
Exp 2. Listening versus Imagining:	
Miles Davis, “So What”	Jazz
Mussorgsky, “Great Gate of Kiev”	Classical
Jukka-Pekka Saraste, Toronto Symphony	
Prokofiev, “Green meadow theme” excerpt from “Peter and the Wolf.” Mario Rossi, Vienna State Opera Orchestra.	Classical
Prokofiev, “the cat theme” excerpt from “Peter and the Wolf.” Mario Rossi, Vienna State Opera Orchestra.	Classical
Bobby Darin, “Mack the Knife”	Pop
Booker T. and the MGs, “Green Onions”	R&B
Sting, “Moon Over Bourbon Street”	Pop
The Kingmen, “Louie Louie”	Rock
Rolling Stones, “Satisfaction”	Rock
Sting, “Englishman in New York”	Pop
Exp 3. Listening, Different Genres	
Bizet, “L’Arlesienne Suite #1.” Slovak Philharmonic Orchestra, Anthony Bramall	Classical
Prokofiev, “Alexander Nevsky.” Lille National Orchestra, Jean-Claude Casadesus	Classical
Britney Spears, “Baby One More Time”	Hot 100
Avril Lavigne, “Keep Holding On”	Hot 100
Spice Girls, “Wannabe”	Hot 100
i5, “Scream Shout”. (Spice Girls soundalikes, control for unfamiliar Hot 100).	Hot 100
Kristy Frank, “Unattached”. (Avril Lavigne soundalike, control for unfamiliar Hot 100).	Hot 100
Chantelle Paige, “Untouchable.” (Britney Spears soundalike, control for unfamiliar Hot 100).	Hot 100
Frank Sinatra, “Mack the Knife”	Jazz
James Brown, “I Got You (I Feel Good)”	R&B
James Brown, “Papas Got a Brand New Bag”	R&B
Percy Thrillington, “Uncle Albert”	Muzak
Bert Kaempfert, “Blue Moon”	Muzak
Freddie MacGregor, “Across the Border”	Reggae
Piazzolla, “Libertango”. Yo-Yo Ma.	Tango
The Beatles, “Girl”	Rock

4th degree B-spline interpolation. For localizer estimation, the processed functional scans were smoothed with an 8 mm isotropic filter. For rendering of localizer results, the cortical surface was defined on the MPRAGE scan using Freesurfer (<http://surfer.nmr.mgh.harvard.edu>). T-maps were projected to this space using rigid body transformations derived from the mapping of the mean EPI-BOLD scans with the native space MPRAGE. The MPRAGE scan was also spatially normalized by non-linear warping to the MNI atlas in SPM12 so that coordinates of local maxima from the localizer t-map could be described in standard space.

Localizer

The goal of this step was to define a set of voxels (not necessarily co-located in a single, contiguous region of interest) that were active during the perception of music relative to a silent baseline. This voxel set would then be used for all subsequent RDA analyses. To do this, a design matrix was created with the 16 trials in experiment 3 (listening to familiar songs in different styles) as a listening condition and the 5 rest scans of experiments 1 and 2 were included as a silent rest condition. A contrast between listening and silence conditions was defined and all voxels passing a significance threshold of $p < 0.001$ and comprising a cluster ≥ 10 voxels were included. The regions included in this analysis are shown in Figure 1, and the local maxima itemized in Table 2. We note that this is a localizer for this individual, and we make no claims that the

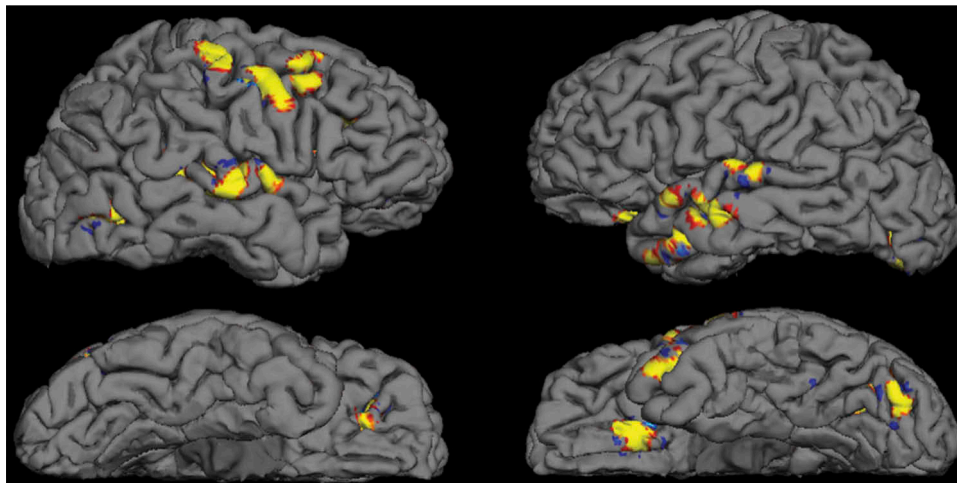


Figure 1. Localization of music related listening versus silence. Significant differences of fMRI BOLD activity during listening to 16 familiar songs are compared to a set of quiet resting conditions ($p < 0.001$, cluster > 10 voxels). Results are projected onto the surface of the participant's own structural MRI scan for visualization.

Table 2. Localization of music listening activity.

MNI Coordinates			Hemisphere	Anatomic Location
X	Y	Z		
38	-14	60	Right	Middle Frontal Gyrus
42	18	52	Right	Middle Frontal Gyrus
56	-2	52	Right	Central Sulcus
44	-28	24	Right	Insula
50	-22	0	Right	Planum Temporale
64	-18	0	Right	Planum Temporale
60	-4	0	Right	Superior Temporal Gyrus
44	-72	0	Right	Posterior Fusiform/Occipital
28	40	-10	Right	Orbital Frontal Cortex
-64	-18	8	Left	Planum Temporale
-30	-76	-4	Left	Posterior Fusiform/Occipital
-60	-8	-16	Left	Anterior Middle Temporal Gyrus
-52	6	-30	Left	Anterior Superior Temporal Sulcus
-30	-62	-42	Left	Cerebellum
-10	26	-26	Left	Orbital Frontal Cortex

results are normative or generalizable to other subjects or sessions for this individual. Instead, the goal of this step is to find a set of voxels that are active when this case study was listening to music relative to silence. All subsequent analyses are evaluated within this set of 527 voxels.

Encoding analysis

Each of the three experiments was modeled within the general linear framework using unsmoothed data in native space. For each experiment, each condition was estimated as a beta image. For experiment 1, the four repeated trials of rest, global composition, lyrics, melody, rhythm, prose, and paint were modeled as 6 separate conditions in the design matrix (with the rest condition serving as "baseline"). For experiment 2, there were 10 different songs, each imagined, and then listened to by the subject as well as 1 rest scan. The design matrix including these 20 music trials each treated as separate conditions as well as a "baseline" rest condition consisting of the 1 rest trial from experiment 2 and the 4 rest scans from experiment 1. For Experiment 3, the design matrix consisted of 16 conditions (i.e., 16 different songs) where the subject was

listening to music. A rest condition consisting of 5 trials obtained from experiments 1 and 2 was used as the baseline. For each of the 3 experiments, we then extracted the beta values at the voxels within our localizer for every condition. In addition, the residuals from voxels within the localizer each experiment were saved and subsequently used to calculate an unbiased estimate of the covariance between voxels of the beta images. To estimate an invertible covariance matrix we used the `covdiag.m` shrinkage function available in the RSA Toolbox (<http://www.mrc-cbu.cam.ac.uk/methods-and-resources/toolboxes/license/>) (Ledoit & Wolf, 2004; Nili et al., 2014).

Representational dissimilarity analysis

A pattern of activity within a set of voxels can be described as a 1-dimensional vector. Given two of these vectors from different experimental conditions, we seek to use a reliable measure of distance that captures how similar or dissimilar they are. There are a number of strategies for doing this, with clear tradeoffs between analytic approaches and preprocessing choices (Walther et al., 2016). Intuitively, the Euclidean distance might be sufficient to characterize distance. However, voxels that covary strongly across conditions will bias this distance estimate and results will be sensitive to the number of voxels in the vector. Here we use the Mahalanobis distance, which has been shown to be remarkably robust across voxels sets of different sizes (Walther et al., 2016). This is the Euclidean distance, inversely weighted by the unbiased estimate of the covariance (using the residuals of experiments 1–3 as described above). We emphasize that the resulting measures reflect distances between multivoxel patterns of activity and are not statistical comparisons of univariate mean activity within the localizer or network correlations.

Linkage representation

For each experiment, the distance between each pair of multivoxel activity patterns, i.e., between each pairs of conditions, can be represented as a 2-D matrix (Figures 2–4). However, these matrices by themselves can be difficult to evaluate. To

aid in data interpretation, each distance matrix is replotted as a dendrogram using the MATLAB functions `dendrogram.m` and `optimalleaforder.m`.

Results

Experiment 1: cortical representation of music composition

Starting with the voxels we identified in the localizer step, we measured the multivoxel pattern of activity during *de novo* music composition of a complete song ("global" condition). We then tested how this pattern differed when the composer focused attention on just the song melody, song rhythm or combined melody & rhythm. After a rest, the composer was also asked to compose unrelated prose, or to imagine painting a canvas. As shown in Figure 2(a), there are obvious differences in the pairwise pattern distances between each of these conditions. Because this is only a single estimation from a case study, we cannot make statistical claims about which of these differences are more or less significant. Nevertheless, the obvious pattern is worthy of evaluation as an exploratory analysis and consistent with our case study approach. These distances can be visualized as a linkage diagram, shown in Figure 2(b). As an example of how to interpret this figure, consider the path distance in Figure 2(b) required to travel from the Canvas to Prose condition compared to the path

distance to travel from Melody to Rhythm. This linkage diagram demonstrates that the differences of the multivoxel pattern of activity between conditions have a clear structure.

The two-dimensional Mahalanobis distance metric (Figure 2(a)) revealed maximum dissimilarity between the four music tasks and painting. Interestingly, prose was more similar to the music composition tasks than painting was, and prose reached its maximum similarity with rhythm. This raises the intriguing suggestion that the musician treats prose more as oration or poetry than as straightforward speech. The most tightly coupled pair in representational space were melody and melody + rhythm, suggesting that melody is the driving force in this composer's mental imagery – rhythm alone is less similar to melody + rhythm than melody is melody + rhythm. Consistent with music theory (Meyer, 1973) melody and rhythm are tightly coupled.

Experiment 2: cortical representations of imagined and listened familiar songs

In the second experiment, the musician imagined each of 10 familiar songs of different styles. Each imagined song was directly followed by the auditory presentation of the same song.

All of the listen-imagine pairs plotted within each square along the diagonal in Figure 3(a) demonstrate very high similarity (Figure 3(a)), as indicated by the blue color coding. The

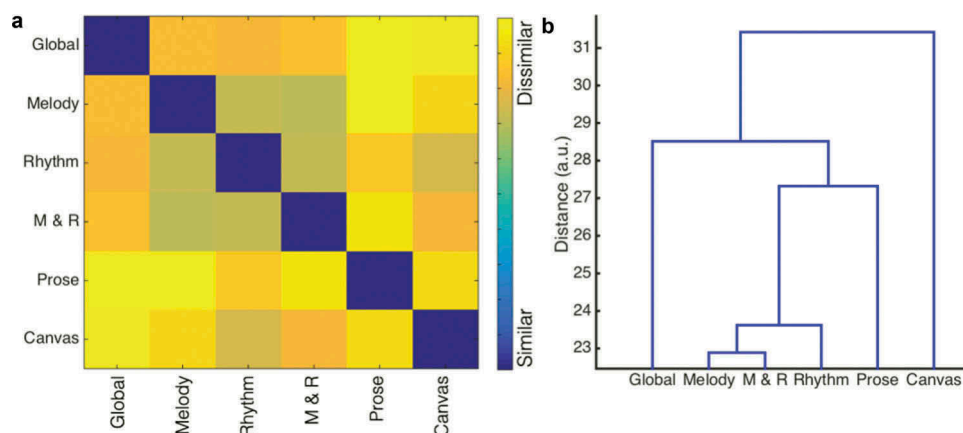


Figure 2. Multivoxel pattern analysis distances during music composition. (a) The Mahalanobis distance between the MVPA of each pair of conditions is summarized (yellow, large distance, blue 0 distance). (b) The same distances replotted as a dendrogram (a.u., arbitrary units). [To view this figure in color, please visit the online version of this Journal.]

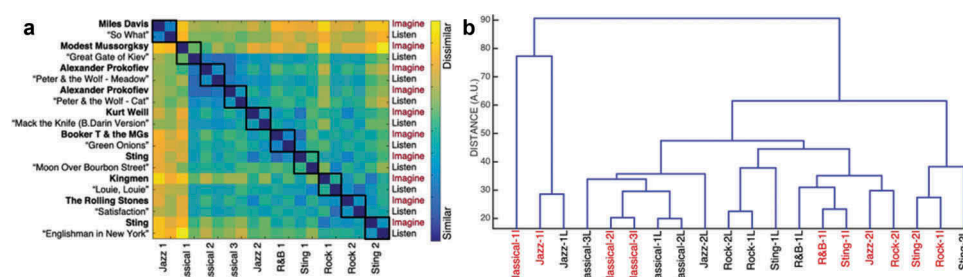


Figure 3. Multivoxel pattern analysis distances during imagined and listened music. (a) The Mahalanobis distance between the MVPA of each pair of conditions is summarized (yellow, large distance, blue 0 distance). The squares surround the imagine-listen pair of each song. (b) The same distances are replotted as a dendrogram (a.u., arbitrary units, imagined in red, listen in back). [To view this figure in color, please visit the online version of this Journal.]

identity diagonal is of course the darkest blue, and in seven of the ten cases, the adjacent square (directly to the right, or directly underneath, representing the listen-imagine comparison) is also a shade of blue. This strongly suggests that the participant's neural activation while imagining music is very close to his neural activation when listening. The selection with the weakest listen-imagine link (light green squares) was from Mussorgsky's "Pictures at an Exhibition," the movement titled "The Great Gate of Kiev." The musician volunteered during debriefing that he mistakenly imagined a different movement of "Pictures" than the one instructed ("Promenade"). The remaining two selections with weaker listen-imagine links than the modal value – although still strong, shown by blue-green squares in the Mahalanobis plot of Figure 3(a) – are The Kingmen's "Louie, Louie" and the musician's own composition, "Moon Over Bourbon Street." One possible explanation for the result with "Louie Louie" is that the song exists in dozens of different versions by different artists, which could have led to some proactive and retroactive interference with the memory trace. One might ask why the participant isn't as familiar with his own song as he is others'. We speculate that, like many musicians, he doesn't listen to his own compositions often, and he performs this particular song less often than "Englishman in New York" and so it is not as fresh in his mind.

With respect to the dendrograms (Figure 3(b)), there are two hypotheses about the ways that the resulting multivoxel patterns might cluster. First, they could cluster in terms of whether or not songs were imagined or listened to. Second, they might cluster in terms of song style. As shown in Figure 3(a), there are obvious pairwise differences in the MVPA Mahalanobis distance between songs (the yellow, gold and light green shades). When replotted as a dendrogram in Figure 3(b), both hypothesized patterns are revealed to be present, with the stronger connections among songs being driven by genre, not by listen versus imagine condition. While there are some local exceptions, there is clear grouping of imagined songs according to genre. The dendrogram divides the entire set into two clusters at the first hierarchical level, with Miles Davis "So What" and the imagined version of Mussorgsky's "The Great Gate of Kiev" clustering together on the left, and all the other pieces clustering on the right. Note

that all the classical pieces are contiguous, both listened and imagined, with the prominent exception of the imagined version of "The Great Gate of Kiev" (first entry on the left), probably for the reason mentioned above, that S imagined a different part of the work than instructed.

We note also that the pattern among the three classical pieces, one by Mussorgsky and two by Prokofiev. Not surprisingly, perhaps, the "Meadow" theme from "Peter & The Wolf" was most similar in its imagined version to the "Cat" theme from "Peter & The Wolf." However, in the listened versions, the "Meadow" theme was found to be more similar to the Mussorgsky piece. This could be because both pieces entail full orchestra (something that would be very salient perceptually) whereas the "Cat" theme is solo bass clarinet.

There are several cases of listened and imagined songs being farther apart than listened songs to one another, suggesting that an imagined representation will not necessarily be identical to the representation formed from direct listening, and that listening and imagining recruit different, though perhaps overlapping, neural regions. Again, we emphasize that these patterns are a single estimation and exploratory in nature.

Note the proximity of Sting's "Moon Over Bourbon Street" with Booker T. & The MG's "Green Onions." Both are in the key of F minor, are at the same tempo (132 bpm), and both have a swing rhythm. Interestingly, the *imagined* version of rock song 2 ("Satisfaction", E maj, 136 bpm) was quite similar to both the listened and imaged version of the musician's other self-composition ("Englishman in New York," B min, 104 bpm). Closer inspection reveals that what is common to these two songs is that the first bass note of each song is the distinctive low (open) E. Our participant is a bass player – it is conceivable that what binds these two songs together in his own idiosyncratic representational space is the first note he would play on his instrument if he were to perform these songs.

Cortical representations for familiar song styles

In the final experiment the musician listened to a set of familiar songs covering a broad range of music styles. In addition, several songs by unfamiliar musicians that were soundalike styles of well known musicians were included.

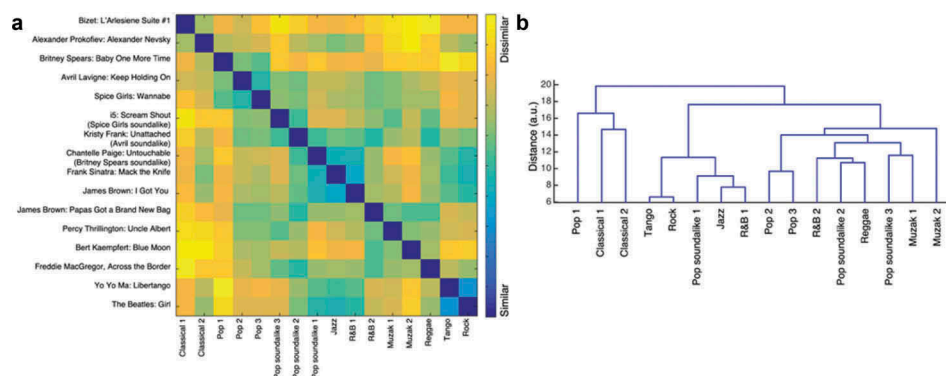


Figure 4. Multivoxel pattern analysis distances during listening to different music styles. (a) The Mahalanobis distance between the MVPA of each pair of conditions is summarized (yellow, large distance, blue 0 distance). (b) The same distances replotted as a dendrogram (a.u., arbitrary units). [To view this figure in color, please visit the online version of this Journal.]

From this, we hypothesized that there might be clustering of cortical activity patterns in terms of either style or familiarity with specific songs. As shown in Figure 4(a), there is remarkable variation in the pair-wise MVPA distance between different songs. Replotting this as a dendrogram in Figure 4(b), there are some important clusterings in terms of overall song style. For example, the two different classical pieces are closely linked, as are the Muzak and unfamiliar pop music. In terms of soundalike songs, there is no clear clustering, suggesting that song familiarity in and of itself is not influencing the pattern distances.

As predicted, the pop songs showed the highest similarity among one another (Figure 4(a)), suggesting a lack of representational distinctiveness in this particular listener. The Spice Girls were perceived as being not very different from Avril Lavigne or Britney Spears. Evidently, pop is pop. The soundalike groups also showed high similarity to one another and the name brand pop. Predictably, maximum dissimilarity was found between classical music and reggae, and between pop and tango.

Why might the Muzak cluster by itself in the dendrogram (Figure 4(b))? A priori, one might imagine that it would be perceived as similar to classical, because of its full orchestration and violins. Or that it would be represented as similar to pop or rock, because the two selections are in fact light symphonic renderings of pop songs (one by Paul McCartney and one by Rogers and Hart). When asked about it, the participant provided an explanation. "That doesn't surprise me at all. I don't see those other genres as being separate. I see quality music as being quality music. And Muzak, which I think is a curse, is something I want to shun... If I hear music that is just expected, then it doesn't interest me at all."

Some of the closest similarity Mahalanobis distances were between the tango (Piazzola's "Libertango") and the rock song (The Beatles' "Girl"). Both are in minor keys (Am and Cm respectively) and both use a similar, prominent three-note motif, from E to F and back to E (Figure 5). A variation – a stretching of that motif – also appears in both compositions, a prominent descent from C to A (in the upper octave of "Libertango," and in the lower octave of "Girl" with a passing tone of B).

Discussion

In this case study of a master musician and composer, we sought to learn about the neuroanatomical substrates of music listening, imagining, composing, and about how different genres of music might be represented cortically. This was not a controlled experiment – there was no null hypothesis or control group – but we structured our study in a way that stimuli could serve as their own controls. The resulting activation patterns we found were interpretable and coherent.

In Experiment 1, we examined creativity in real time by asking the composer to create three new pieces of music in his mind, and we compared that to his creating three new pieces of prose or three paintings in his mind. After creating each song, we asked him to focus on separate elements of the new composition: melody, rhythm, or melody + rhythm combined. As we hypothesized, the lone visual task was differentiable from the audio tasks based on patterns of neural activation in our regions of interest. Melody + rhythm combined showed great similarity to melody alone than to rhythm alone, suggesting that melody is more than half of a song, at least in this composer's mind, and for these particular songs.

In Experiment 2, we examined cortical representations of imagined and listened familiar songs. Our two principal findings were that seven of the ten

listened-imagine pairs showed very high similarity, and that certain songs tapped into similar neural responses, consistent with objective properties of those songs, such as pitch, key, motif, or orchestration. Classical music tended to cluster, in both listened and imagined versions, and for other selections, the listened and imagined versions were separable. Previous studies have shown substantial overlap of neural activation between listening and imagining (Herholz, Halpern, & Zatorre, 2012; Herholz, Lappe, Knief, & Pantev, 2008; Zatorre et al., 1996). Our findings are consistent with these, and suggest that individual selections, in part governed by genre, created even tighter activations in this particular musician, consistent with an emerging body of work about the latent structure of musical preferences (Rentfrow et al., 2011, 2012). Our findings are highly consistent with those of Kleber et al. (2007), in which musicians imagined and sang music. Unlike the



Figure 5. Comparison of the motifs of "Libertango" by Piazzolla and "Girl" by Lennon & McCartney.

previous studies mentioned, which used popular songs, Kleber, et al used classical music (Italian operatic arias) and, like us, found that fMRI patterns of activation during listening and imaging of classical music were very similar.

In Experiment 3, we hypothesized that musical style and genre might underlie representational differences in the musician's brain. We presented music spanning a wide range of styles and genres, including popular (top 100), classic rock, jazz, R&B, tango, Muzak (easy listening), classical, tango, and reggae. We discovered remarkable variation in activation patterns among the different selections. To control for familiarity – the massive exposure that pop music has (it can be hard to avoid it in public spaces) – we presented unknown songs by soundalike artists alongside the well-known artists. The popular songs showed very high similarity to one another, irrespective of familiarity. The Muzak (easy listening) selections, which S later described as “a curse” and something he wants to “shun”, clustered far away from anything else in representational space, indicating his aversion toward them. Additionally, we found great dissimilarity between classical music and reggae, and between pop and tango. Reggae and tango are driven more by rhythm than classical and pop are, and this may explain their representational distinctiveness.

It is well-established that music recruits differentiable neural networks from speech (Abrams et al., 2011; Norman-Haignere, Kanwisher, & McDermott, 2015). Although we did not test speech, the activations we found for music listening were consistent with previous studies, showing bilateral activation of music versus silence in the cortex, as well as activation in the cerebellum (Blood & Zatorre, 2001; Janata & Grafton, 2003; Levitin & Tirovolas, 2009; Figure 1). Notably, we found activation in temporal lobe regions, planum temporale, orbital frontal cortex, insula, and posterior fusiform/occipital regions.

We caution against generalizing these findings to other musicians. These data are obviously based on an *n* of one, and we make no claims about what might be found in other musicians and composers. Nevertheless, an important contribution of the work we report here is the introduction of a set of analysis tools that can be used by investigators in future studies of music and expertise. Indeed, given a suitable task, the Mahalanobis distance metric and dendrograms could be applied to BOLD response data in studying other areas of expertise, such as athletics, chess, reasoning, and visual art.

Acknowledgements

Our thanks to the technicians who assisted with acquisition of the brain images, André Cormier and Mike Ferreira, to Tracy Bufferd, Kathy Schenker, and William Francis for logistical help, and to Sting for generously granting his time and insights to this study. We also thank Jamshed Bharucha for assistance with the study design, David Colman for access to the scanner, and Keith Worsley for assistance with preliminary analyses.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Sciences and Engineering Research Council of Canada [228175-10] to the first author.

ORCID

Daniel J. Levitin  <http://orcid.org/0000-0002-9534-1494>

Scott T. Grafton  <http://orcid.org/0000-0003-4015-3151>

References

- Abrams, D. A., Bhatara, A., Ryali, S., Balaban, E., Levitin, D. J., & Menon, V. (2011). Decoding temporal structure in music and speech relies on shared brain resources but elicits different fine-scale spatial patterns. *Cerebral Cortex*, 21, 1507–1518. doi:10.1093/cercor/bhq198
- Attneave, F., & Curlee, T. E. (1977). Cartesian organization in the immediate reproduction of spatial patterns. *Bulletin of the Psychonomic Society*, 10, 469–470. doi:10.3758/BF03337700
- Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences*, 98, 11818–11823. doi:10.1073/pnas.191355898
- Boring, E. (1923). *A history of experimental psychology*. New York: Pendragon.
- Brown, D. (1991). *Human universals*. New York: McGraw-Hill.
- Bunzeck, N., Wuestenberg, T., Lutz, K., Heinze, H.-J., & Jancke, L. (2005). Scanning silence: Mental imagery of complex sounds. *Neuroimage*, 26, 1119–1127. doi:10.1016/j.neuroimage.2005.03.013
- Cadlin, A., McAdams, S., Smith, B. K. & Winsberg, S. (2005). Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones. *The Journal of the Acoustical Society of America*, 118, 471–482. doi:10.1121/1.1929229
- De Valois, R. L., & De Valois, K. D. (1997). Neural coding of color. In A. Byrne & D. R. Hilbert (Eds), *Readings on color*, Vol. 2: *The science of color* (pp. 93–140). Cambridge, MA, US: The MIT Press.
- Halpern, A. R., & Zatorre, R. J. (1999). When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex*, 9, 697–704. doi:10.1093/cercor/9.7.697
- Herholz, S., Lappe, C., Knief, A., & Pantev, C. (2008). Neural basis of music imager and the effect of musical expertise. *European Journal of Neuroscience*, 28, 2352–2360. doi:10.1111/j.1460-9568.2008.06515.x
- Herholz, S. C., Halpern, A. R., & Zatorre, R. J. (2012). Neuronal correlates of perception, imagery, and memory for familiar tunes. *Journal of Cognitive Neuroscience*, 24, 1382–1397. doi:10.1162/jocn_a_00216
- Hongwen, K., Hebert, M., Efros, A. A., & Kanade, T. (2015). Data-driven objectness. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37, 189–195. doi:10.1109/TPAMI.2014.2315811
- Houlihan, K., & Levitin, D. J. (2011, August). *Recognition of melodies from rhythm and pitch*. Paper presented at the 2011 bi-annual meeting of the Society for Music Perception and Cognition (SMPC), Rochester, NY.
- Janata, P., & Grafton, S. T. (2003). Swinging in the brain: Shared neural substrates for behaviors related to sequencing and music. *Nature Neuroscience*, 6, 682–687. doi:10.1038/nn1081
- Jaśkowski, P. (1999). Reaction time and temporal-order judgment as measures of perceptual latency: The problem of dissociations. *Advances in Psychology*, 129, 265–282.
- Keefe, S. P. (2006). Rochlitz, (Johann) Friedrich. In C. Eisen & S. P. Keefe (Eds.), *The Cambridge mozart encyclopedia*. Cambridge: Cambridge University Press.
- Kleber, B., Birbaumer, N., Veit, R., Trevorrow, T., & Lotze, M. (2007). Overt and imagined singing of an Italian aria. *NeuroImage*, 36, 889–900. doi:10.1016/j.neuroimage.2007.02.053
- Konrad, U. (1992). *Mozarts Schaffensweise: Studien zu den Werkautographen, Skizzen und Entwürfen*. Göttingen: Vandenhoeck & Ruprecht.

- Kraemer, D. J., Macrae, C. N., Green, A. E., & Kelley, W. M. (2005). Musical imagery: Sound of silence activates auditory cortex. *Nature*, 434, 158–158. doi:10.1038/434158a
- Krumhansl, C. L. (2010). Plink: "Thin slices" of music. *Music Perception*, 27, 337–354. doi:10.1525/mp.2010.27.5.337
- Ledoit, O., & Wolf, M. (2004). Honey, I shrunk the sample covariance matrix. *The Journal of Portfolio Management*, 30, 110–119. doi:10.3905/jpm.2004.110
- Levitin, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception & Psychophysics*, 56, 414–423. doi:10.3758/BF03206733
- Levitin, D. J. (1999). Memory for musical attributes. In P. R. Cook (Ed.), *Music, cognition and computerized sound: An introduction to psychoacoustics* (pp. 209–227). Cambridge, MA: The M.I.T. Press.
- Levitin, D. J. (2007). *This is your brain on music*. New York: Dutton/Penguin Random House.
- Levitin, D. J. (2009). *The world in six songs: How the musical brain created human nature*. New York: Dutton/Penguin Random House.
- Levitin, D. J. (2012). What does it mean to be musical? *Neuron*, 73, 633–637. doi:10.1016/j.neuron.2012.01.017.
- Levitin, D. J., Chordia, P., & Menon, V. (2012). Musical rhythm spectra from Bach to Joplin obey a 1/f power law. *Proceedings of the National Academy of Sciences*, 109, 3716–3720. doi:10.1073/pnas.1113828109
- Levitin, D. J., & Cook, P. R. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute. *Perception & Psychophysics*, 58, 927–935. doi:10.3758/BF03205494
- Levitin, D. J., & Tirovolas, A. K. (2009). Current advances in the cognitive neuroscience of music. *Annals of the New York Academy of Sciences*, 1156, 211–231. doi:10.1111/j.1749-6632.2009.04417.x
- Meyer, L. B. (1973). *Explaining music: Essays and explorations*. Berkeley, CA: University of California Press.
- Nili, H., Wingfield, C., Walther, A., Su, L., Marslen-Wilson, W., & Kriegeskorte, N. (2014). A toolbox for representational similarity analysis. *PLoS Computational Biology*, 10, e1003553. doi:10.1371/journal.pcbi.1003553
- Norman-Haignere, S., Kanwisher, N. G., & McDermott, J. H. (2015). Distinct cortical pathways for music and speech revealed by hypothesis-free voxel decomposition. *Neuron*, 88, 1281–1296. doi:10.1016/j.neuron.2015.11.035
- Pochmursky, C. (Writer/Director) (2009, January 31). *The musical brain* [Television broadcast]. In V. Dylun (Producer), Mississauga, ON: Matter of Fact Media Inc. Production in association with CTV Television Inc. and National Geographic Channel.
- Porter, N. (1871). *The human intellect*. New York, NY: Scribner's.
- Rentfrow, P. J., Goldberg, L. R., & Levitin, D. J. (2011). The structure of musical preferences: A five-factor model. *Journal of Personality and Social Psychology*, 100, 1139–1157. doi:10.1037/a0022406
- Rentfrow, P. J., Goldberg, L. R., Stillwell, D. J., Kosinski, M., Gosling, S. D., & Levitin, D. J. (2012). The song remains the same: A replication and extension of the MUSIC model. *Music Perception: An Interdisciplinary Journal*, 30, 161–185. doi:10.1525/mp.2012.30.2.161
- Schellenberg, E. G., Iverson, P., & McKinnon, M. C. (1999). Name that tune: Identifying popular recordings from brief excerpts. *Psychonomic Bulletin & Review*, 6, 641–646. doi:10.3758/BF03212973
- Shepard, R. N. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review*, 91, 417–447. doi:10.1037/0033-295X.91.4.417
- Shepard, R. N., & Cooper, L. A. (1992). Representation of colors in the blind, color-blind, and normally sighted. *Psychological Science*, 3, 97–104. doi:10.1111/psci.1992.3.issue-2
- Solomon, M. (1995). *Mozart: A life*. New York: Harper Perennial.
- Tolman, E. (1926). Habit formation and higher mental processes in animals. *Psychological Bulletin*, 25, 24–53. doi:10.1037/h0075466
- Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the "internal clock". *Psychological Monographs: General and Applied*, 77, 1–31. doi:10.1037/h0093864
- Tversky, B. (1981). Distortions in memory for maps. *Cognitive Psychology*, 13, 407–433. doi:10.1016/0010-0285(81)90016-5
- Vigeland, C. (2009). *The mostly mozart guide to Mozart*. New York: Wiley.
- Walther, A., Nili, H., Ejaz, N., Alink, A., Kriegeskorte, N., & Diedrichsen, J. (2016). Reliability of dissimilarity measures for multi-voxel pattern analysis. *NeuroImage*, 137, 188–200.
- White, B. W. (1960). Recognition of distorted melodies. *The American Journal of Psychology*, 73, 100–107. doi:10.2307/1419120
- Yoo, S. S., Lee, C. U., & Choi, B. G. (2001). Human brain mapping of auditory imagery: Event-related functional MRI study. *NeuroReport*, 12, 3045–3049. doi:10.1097/00001756-200110080-00013
- Zaentz, S. (Producer), & Forman, M. (Director). (1984). *Amadeus* [Motion picture]. Los Angeles, CA: Orion Pictures.
- Zatorre, R. J., & Halpern, A. R. (2005). Mental concerts: Musical imagery and auditory cortex. *Neuron*, 47, 9–12. doi:10.1016/j.neuron.2005.06.013
- Zatorre, R. J., Halpern, A. R., Perry, D. W., Meyer, E., & Evans, A. C. (1996). Hearing in the mind's ear: A PET investigation of musical imagery and perception. *Journal of Cognitive Neuroscience*, 8, 29–46. doi:10.1162/jocn.1996.8.1.29
- Zhang, X., Furtlehner, C., & Sebag, M. (2008). Data streaming with affinity propagation. In W. Daelemans & K. Morik (Eds.), *Machine learning and knowledge discovery in databases* (pp. 628–643). Berlin: Springer.