scientific reports

Check for updates

OPEN

An ALE meta-analytic review of top-down and bottom-up processing of music in the brain

Victor Pando-Naude[1,3]✉, Agata Patyczek[2,3], Leonardo Bonetti[1] & Peter Vuust[1]

A remarkable feature of the human brain is its ability to integrate information from the environment with internally generated content. The integration of top-down and bottom-up processes during complex multi-modal human activities, however, is yet to be fully understood. Music provides an excellent model for understanding this since music listening leads to the urge to move, and music making entails both playing and listening at the same time (i.e., audio-motor coupling). Here, we conducted activation likelihood estimation (ALE) meta-analyses of 130 neuroimaging studies of music perception, production and imagery, with 2660 foci, 139 experiments, and 2516 participants. We found that music perception and production rely on auditory cortices and sensorimotor cortices, while music imagery recruits distinct parietal regions. This indicates that the brain requires different structures to process similar information which is made available either by an interaction with the environment (i.e., bottom-up) or by internally generated content (i.e., top-down).

Music is a highly multifaceted aspect of the human experience and therefore a privileged tool to gain insights into the key features of the human brain, its structure, and supporting neural processes. One of such key features is the ability to constantly integrate information and coordinate a variety of different tasks. Indeed, a successful existence depends on the brain's capacity to combine external stimuli with internal representations to ultimately select the best response to the ever-changing environment[1]. The processes which regulate such needs have often been framed in the notorious "top-down/bottom-up" dichotomy. This simplistic approach has been redefined by the predictive coding theory (PC) which intends to set a framework that explains the complexity of the flow of information in the brain. Within this framework, music is an ideal example of the constant integration of multi-modal top-down and bottom-up information, as it includes different processes such as perception, imagery, and production/creation[2]. Despite decades of research in the field of music cognition, there is no comprehensive study investigating brain activation associated to these three distinct modalities of music processing.

Throughout the years, key examples of both top-down and bottom-up processes have been revealed. On the one hand, it has been shown that environmental stimuli activate different brain cortices, coherently with the chosen sensory modality[3] (e.g., visual inputs activate visual brain regions, while auditory inputs activate auditory cortices). According to the PC, bottom-up processing is represented as a series of loops formed by ascending and descending connections between neighbouring levels in a hierarchically organized system (feedforward-feedback). This system is reliable and fast, with a semi-hardwired architecture sufficiently stable to allow for rapid processing of stimuli, and that can be dynamically re-shaped. On the other hand, top-down mechanisms have been described in diverse manners including anatomical (connections between levels of hierarchy), cognitive (hypothesis-driven), *gestaltist* (modulation of bottom-up), and dynamic (entrainment of neuronal populations by oscillatory activity)[4]; all suggesting that the system allows individuals to predict future events and stimuli. Notably, the term top-down should always be challenged when used[5]. According to the PC, top-down processes can be seen as influences that allow the semi-hardwired network to be flexible and ensure efficient and reliable sensory processing. In other words, the system takes as much new information as possible (perception) to construct predictions about the environment to rapidly apply it in a behaviour (action). Taking this in consideration, it is clear that bottom-up and top-down mechanisms are not opposites, and that ascending and descending connections are involved in both processes.

In this context, music is an ideal domain to study the constant integration of multi-modal mechanisms[6], as music engages audio-motor coupling and other cognitive and emotional mechanisms such as pleasure[7]. Thus,

[1]Center for Music in the Brain, Department of Clinical Medicine, Aarhus University and The Royal Academy of Music Aarhus/Aalborg, Universitetsbyen, 3-0-17, 8000 Aarhus C, Denmark. [2]MR Center of Excellence, Center for Medical Physics and Biomedical Engineering, Medical University of Vienna, Vienna, Austria. [3]These authors contributed equally: Victor Pando-Naude and Agata Patyczek. ✉email: pandonaude@clin.au.dk

how the brain processes music may represent a privileged tool to gain insights into such central topics. Music is built upon several concurrent bottom-up processes such as perception of sounds and rhythms. Moreover, it also comprises a number of top-down mechanisms such as music production and music imagery, the latter being the abstract mental manipulation of musical sounds[8]. On the neural level, previous studies investigating such processes have provided a wide array of underlying active brain areas, spanning from auditory and sensorimotor cortices[9] to cerebellum[2], cingulum[10], basal ganglia[11], hippocampus[12], and amygdala[13]. In light of this, music may offer an ideal opportunity to investigate similarities and differences between top-down and bottom-up processes in complex multi-modal brain functioning.

A large number of high-quality studies on music neuroscience already exists. However, individual studies do carry the limitation of small sample size and therefore consequent low reliability[14]. Recent advancements in statistical methods and meta-analyses allow us to overcome restrictions related to data protection regulation[15], gaining insights from a large population size and embrace "big data" human imaging. Currently, there are a number of coordinate-based algorithms for meta-analysis of neuroimaging studies such as the Activation Likelihood Estimation (ALE)[16] and kernel density analysis (KDA)[17]. The ALE method reports location probabilities associated with each foci (set of x, y, and z coordinates), whereas KDA shows the number of foci surrounding a specific voxel.

In the field of music cognition, the ALE methodology has been used to investigate several independent features of music cognition including recruitment of motor areas[2], auditory-motor entrainment[18], hierarchical auditory organisation[19], and music-evoked emotions[20]. Further, projects following open science principles such as the BrainMap[21] platform and its GingerAle[14] meta-analysis algorithm enable sharing of data and so accelerate progress in human brain mapping[22]. These factors converge to make the BrainMap ALE meta-analysis a convenient and appropriate method to compare brain activation across studies of music cognition.

Thus, for the first time, we have performed coordinate-based meta-analyses (CBMA) of a wide range of functional magnetic resonance imaging (fMRI) studies using state-of-the-art methods such as ALE, aiming to assess the neural mechanisms underlying music cognition. As a secondary aim, we sought to assess the broad nature of musical processing as a model of cognition itself. Specifically, appraising perception, production, and imagery in light of top-down and bottom-up processes and the recent theory of predictive coding. In this way, we explore the common and distinct patterns of brain activation during music perception, production, and imagery; and discuss how the brain activity associated with the three modalities may reflect the top-down and bottom-up nature of the brain.

Results

A total of 1707 articles were identified through database searching, and after removing 497 duplicates, 1210 articles were initially screened by title and abstract. After excluding 907 records, 303 articles were assessed for eligibility in the full-text screening stage. From these, 130 fulfilled criteria and were included in both qualitative and quantitative synthesis (ALE meta-analysis) (Supplementary Figure 1).

Characteristics of studies. The characteristics of all studies included in the final qualitative and quantitative synthesis of the meta-analysis are shown in Table 1. In music perception, a total of 105 studies and 2035 participants (122 female; age = 26.6 ± 6.7 years) was identified. Musicians were reported in 12% of the included studies, non-musicians in 69% and both in 13%. Neuroimaging method comprised of both fMRI (96%) and PET (4%). Musical features varied across studies including emotion (24%), melody (11%), harmony (9%), timbre (4%), tonality (4%), memory (3%), pitch (3%), structure (3%), rhythm (2%), tension (2%), creativity (2%), while 15% of studies did not specified the musical feature. Most of the auditory stimuli included unfamiliar stimuli (76%).

In music production, a total of 19 studies and 292 participants (122 female; age = 29.8 ± 6.1 years) was identified. Musicians were reported in 63% of the included studies, non-musicians in 16%, and both in 21%. Neuroimaging method comprised of both fMRI (95%) and PET (5%). Musical features varied across studies including harmony (11%), melody (5%), pitch (10%) and rhythm, metre and/or beat (11%). Overall, most of the studied focused on learnt abstracts (68%) compared to creative improvisation (47%). Most of the tasks included unfamiliar content (53%).

In music imagery, a total of 15 studies and 198 participants (108 female; age = 25.5 ± 3 years) was identified. Musicians were reported in 40% of the included studies, non-musicians in 47%, and both in 13%. Neuroimaging method comprised of both fMRI (93%) and PET (7%). Musical features varied across studies including melody (27%), pitch (13%), timbre (7%) and non-specific (53%) of which most included unfamiliar imagery (60%).

MRI quality. MRI quality of the included studies in the meta-analysis was assessed by a set of guidelines for the standardized reporting of MRI studies[23,24]. All studies reported their MRI design, software package and image acquisition, processing and analyses. Overall, good MRI practices were performed in the included studies (Supplementary Table 1 and Supplementary Table 2). Neuroimaging data was acquired in either 1.5 T (27%), 3 T (66%), while 7% of the studies did not specified the magnetic field strength. Most of the studies used the Siemens MRI scanner (50%), others include General Electric (28%), Phillips Electronics (4%), Hitachi (1%), Brucker (4%), Magnex Eclipse (1%), CTI (1%), and few failed to report the scanner (2%). Most of the structural images were acquired using T1-weighted sequence (42%), magnetization-prepared rapid acquisition with gradient echo sequence (MPRAGE) (27%), spoiled gradient recalled acquisition in steady state (SPGR) (9%) with 1mm3-voxel size in 39% of the studies. T2- functional images were acquired mainly using echo-planar imaging (EPI) sequence (86%). Statistical analysis was conducted in either SPM (60%), AFNI (6%), FSL (7%), LIPSIA (5%) BrainVoyager (6%), MATLAB (2%), while 13% of studies did not specify the analysis software.

	Author	Year	Domain	n	Method	Feature	Familiarity	Males	Females	Total	Musicians	Age		Education (years)		Onset (years)		Training (hr/week)	
												Mean	± SD	Mean	± SD	Mean	± SD	Mean	± SD
1	Agustus	2018	Perception	20	fMRI	Melody	Y	10	10	20	Y	66.6	–	–	–	–	–	–	–
2	Alluri	2012	Perception	11	fMRI	Timbre, key, rhythm	N	6	5	11	Y	23.2	3.7	16.1	6	9.1	3.4	2.5	1.2
3	Alluri	2013	Perception	11	fMRI	Non-specific	Y	7	4	11	N	26.4	5	–	–	–	–	–	–
4	Alonso	2016	Perception	22	fMRI	Melody	N	11	11	22	N	24.09	5.45	–	–	–	–	–	–
5	Altten-müller	2014	Production	18	fMRI	Emotion +	N	14	4	18	N	28.7	8.7	8.4	5.7	–	–	–	–
6	Angulo-Perkins	2014	Perception	53	fMRI	Non-specific	N	29	24	53	B	28	8	–	–	–	–	–	–
7	Armony	2015	Imagery	47	fMRI	Non-specific	N	27	20	47	B	26.4	–	–	–	–	–	–	–
8	Bangert	2006	Imagery	14	fMRI	Non-specific	N	6	8	14	Y	–	–	–	–	–	–	–	–
9	Barrett	2016	Perception	12	fMRI	Emotion +	Y	4	8	12	N	22.4	–	–	–	–	–	–	–
10	Barrett	2018	Perception	25	fMRI	Tonality	Y	–	–	–	N	–	–	–	–	–	–	–	–
11	Barrett	2020	Perception	1	fMRI	Creativity	N	0	1	1	Y	44	–	–	–	–	–	–	–
12	Bastepe-Gray	2020	Perception	7	fMRI	Non-specific	Y	7	0	7	Y	22.8	2.1	6.3	1.5	–	–	10	2.5
13	Baumann	2007	Imagery	7	fMRI	Non-specific	N	1	6	7	B	25.7	3.2	0	0	–	–	–	–
14	Bengtsson	2006	Imagery	11	fMRI	Rhythm and melody	Y	11	0	11	Y	33	–	–	–	5.7	1.4	–	–
15	Bengtsson	2007	Perception	11	fMRI	Structure/creativity	Y	11	0	11	Y	32.6	6	–	–	5.7	1.4	–	–
16	Bengtsson	2009	Perception	17	fMRI	Rhythm + beat + metre	N	14	3	17	N	23.6	–	–	–	–	–	–	–
17	Bianco	2016	Perception	29	fMRI	Harmony	N	12	17	29	Y	24.7	2.9	17.2	4.8	7.3	3.08	–	–
18	Bishop	2013	Perception	12	fMRI	Rhythm + beat + metre	Y	6	6	12	N	21.2	3	–	–	–	–	–	–
19	Blood	1999	Perception	10	PET	Emotion +	N	5	5	10	N	–	–	–	–	–	–	–	–
20	Bodner	2001	Perception	7	fMRI	Rhythm + beat + metre	Y	7	–	–	N	–	–	–	–	–	–	–	–
21	Bogert	2016	Production	56	fMRI	Emotion – non-specific	N	–	–	–	N	28.2	8.21	5	–	–	–	2	–
22	Brattico	2016	Production	29	fMRI	Emotion +	Y	14	15	29	B	23.9	3.1	–	–	–	–	–	–
23	Bravo	2017a	Perception	12	fMRI	Harmony	N	5	7	12	N	29	5.16	–	–	–	–	–	–
24	Bravo	2017b	Imagery + Pro-duction	14	fMRI	Harmony	N	7	7	14	N	30.45	3.05	–	–	–	–	–	–
25	Bravo	2020	Perception	16	fMRI	Harmony	N	8	8	16	N	29	5.16	–	–	–	–	–	–
26	Brown	2004	Perception	10	PET	Non-specific	N	5	5	10	N	33.8	–	–	–	–	–	–	–
27	Brown	2007	Perception	11	fMRI	Melody	N	5	6	11	N	24.6	–	–	–	12.3	–	–	–
28	Chapin	2010	Perception	14	fMRI	Emotion – non-specific	N	5	9	14	N	–	–	–	–	–	–	–	–
29	Chen	2008a	Perception	12	fMRI	Rhythm + beat + metre	N	–	–	12	N	24	–	–	–	–	–	–	–
30	Chen	2008b	Perception	24	fMRI	Rhythm + beat + metre	N	12	12	24	N	23.83	–	–	–	–	–	–	–
31	Chiang	2018	Perception	20	fMRI	Structure – normal	N	12	8	20	Y	–	–	–	–	–	–	–	–
32	Danielsen	2014	Perception	19	fMRI	Rhythm + beat + metre	N	11	8	19	N	35.2	6.2	–	–	–	–	–	–
33	Demorest	2009	Production	16	fMRI	Memory	N	8	8	16	N	28.6	–	–	–	–	–	–	–
34	Donnay	2014	Perception	11	fMRI	Harmony	N	11	0	11	Y	38.8	11	–	–	–		–	–
35	Engel	2011	Perception	22	fMRI	Creativity	N	22	0	0	Y	–	–	12.8	6.8	–	–	7.4	5.8
36	Escoffier	2013	Perception	16	fMRI	Emotion – non-specific	N	9	7	16	N	21.7	1.9	–	–	–	–	–	–
37	Fedorenko	2012	Production	12	fMRI	Structure – normal	N	6	6	12	N	–	–	–	–	–	–	–	–
38	Flores-Gutiérrez	2007	Perception	19	fMRI	Emotion + and –	N	11	8	19	N	25	3.05	–	–	–	–	–	–
39	Fujisawa	2011	Perception	12	fMRI	Harmony	Y	12	0	12	N	–	–	–	–	–	–	–	–
40	González-García	2016	Perception	11	fMRI	Harmony	Y	0	11	11	Y	–	–	–	–	–	–	–	–
41	Grahn	2007	Perception	27	fMRI	Rhythm + beat + metre	N	8	19	27	B	24.5	–	–	–	–	–	–	–
42	Grahn	2009a	Perception	35	fMRI	Rhythm + beat + metre	N	23	12	35	N	29.9	7.2	–	–	–	–	–	–
43	Grahn	2009b	Perception	36	fMRI	Rhythm + beat + metre	N	21	15	36	N	29	–	–	–	–	–	–	–
44	Grahn	2013	Perception	24	fMRI	Rhythm + beat + metre	N	11	13	24	B	27	–	–	–	–	–	–	–
45	Green	2008	Perception	21	fMRI	Harmony	N	–	–	–	N	27.3	–	–	–	–	–	–	–
46	Green	2012	Perception	21	fMRI	Familiarity	N	9	12	21	N	27.3	–	–	–	–	–	–	–
47	Green	2018	Perception	22	fMRI	Familiarity	N	17	5	22	N	–	–	–	–	–	–	–	–
48	Halpern	1999	Perception	8	fMRI	Melody	N	3	5	8	B	–	–	–	–	–	–	–	–
49	Halpern	2004	Perception	10	fMRI	Timbre	N	5	5	10	Y	–	–	5	–	–	–	–	–
50	Herdener	2014	Perception	22	fMRI	Rhythm + beat + metre	N	–	–	22	B	–	–	5	–	–	–	–	–
51	Herholtz	2012	Perception	10	fMRI	Non-specific	Y	5	5	10	N	27	–	–	–	–	–	–	–
52	Huang	2016	Perception	18	fMRI	Non-specific	Y	18	0	18	N	–	–	–	–	–	–	–	–
53	Janata	2002a	Perception	8	fMRI	Tonality	Y	–	–	8	N	–	–	–	–	–	–	–	–
54	Janata	2002b	Perception	12	fMRI	Melody	N	5	7	12	Y	28.5	–	8.3	3.5	–	–	–	–
55	Janata	2009	Production	13	fMRI	Memory	N	2	11	13	Y	20	–	–	–	–	–	–	
56	Jeong	2011	Perception	15	fMRI	Emotion + and –	N	5	10	15	N	22.8	3.4	–	–	–	–	–	–
57	Jungblut	2012	Perception	30	fMRI	Rhythm + beat + metre	N	17	13	30	N	26.3	–	–	–	–	–	–	–
58	Khalfa	2005	Perception	13	fMRI	Emotion non-specific	N	8	5	13	N	28	–	–	–	–	–	–	–
59	Kleber	2007	Perception	16	fMRI	Non-specific	Y	5	11	16	Y	31.06	8.27	14.06	7.59	–	–	25.73	7.26
60	Kleider-Offutt	2019	Perception	28	fMRI	Non-specific	N	11	17	28	N	21.4	–	–	–	–	–	–	–

Continued

	Author	Year	Domain	n	Method	Feature	Familiarity	Males	Females	Total	Musicians	Age		Education (years)		Onset (years)		Training (hr/week)	
												Mean	± SD	Mean	± SD	Mean	± SD	Mean	± SD
61	Koelsch	2003	Perception	11	fMRI	Emotion + and −	N	5	6	11	N	24.6	–	–	–	–	–	–	–
62	Koelsch	2014	Perception	20	fMRI	Emotion +	N	10	10	20	N	25.55	4.8	–	–	–	–	–	–
63	Koelsch	2018	Imagery	24	fMRI	Emotion + and −	Y	12	12	24	N	23.39	3.3	–	–	–	–	–	–
64	Kornysheva	2010	Perception	18	fMRI	Rhythm + beat + metre	N	17	11	18	N	25.1	–	–	–	–	–	–	–
65	Langheim	2002	Perception	6	fMRI	Non-specific	Y	2	4	6	Y	27	–	–	–	–	–	–	–
66	Leaver	2009	Production	20	fMRI	Non-specific	N	10	10	20	N	24.7	–	–	–	–	–	–	–
67	Lee	2011	Production + perception	12	fMRI	Melody	N	7	5	12	N	20.4	–	5.7	–	–	–	–	–
68	Lehne	2013	Perception	25	fMRI	Tension, loudness	N	22	13	25	N	23.9	3.6	6.4	3.6	–	–	–	–
69	Levitin	2005	Perception	13	fMRI	Structure	N	6	7	13	Y	–	–	–	–	–	–	–	–
70	Levitin	2016	Production + perception	1	fMRI	Non-specific	Y	1	0	1	Y	55	–	–	–	–	–	–	–
71	Li	2019	Production	31	fMRI	Rhythm + beat + metre	N	15	16	31	N	–	–	7	–	–	–	–	–
72	Limb	2006	Imagery	24	fMRI	Rhythm + beat + metre	N	18	6	24	B	–		–	–	–	–	–	–
73	Limb	2008	Production	6	fMRI	Creativity	N	6	0	6	B	34.2	10.4	–	–	–	–	–	–
74	Liu	2018	Production	12	fMRI	Creativity	N	12	0	12	B	30.3	–	9.8	4.3	–	–	–	–
75	Matthews	2020	Perception	51	fMRI	Rhythm + beat + metre	N	30	21	51	B	–	–	–	–	–	–	–	–
76	Meister	2004	Imagery + Perception	12	fMRI	Learnt	Y	2	10	12	Y	26.6	–	–	–	–	–	–	–
77	Merrill	2012	Imagery	21	fMRI	Non-specific	N	14	7	21	N	24.2	2.4	–	–	–	–	–	–
78	Mizuno	2007	Imagery	18	fMRI	Tonality	N	10	8	10	N	23.1	2.8	12.4	5.8	–	–	–	–
79	Montag	2011	Perception	33	fMRI	Emotion +	Y	6	27	33	N	23.55	5.04	–	–	–	–	–	–
80	Morrison	2003	Perception	6	fMRI	Non-specific	N	2	4	6	B	38.3	–	–	–	–	–	–	–
81	Mueller	2015	Perception	23	fMRI	Emotion +	N	10	13	23	N	25.9	2.9	–	–	–	–	–	–
82	Ohnishi	2001	Perception	28	fMRI	Non-specific	N	14	14	28	B	–	–	–	–	–	–	–	–
83	Park	2013	Perception	12	fMRI	Emotion +	N	5	7	12	Y	20.33	2.14	13.2	3.6	–	–	–	–
84	Park	2014	Production	24	fMRI	Emotion −	N	20	14	24	B	–	–	–	–	–	–	–	–
85	Parsons	2005	Perception	8	PET	Learnt	Y	3	5	8	Y	–	–	–	–	–	–	–	–
86	Pereira	2011	Imagery	27	fMRI	Familiarity	N	9	18	27	N	32	–	–	–	–	–	–	–
87	Peretz	2009	Imagery	9	fMRI	Familiarity	Y	0	9	9	N	–	–	–	–	–	–	–	–
88	Petrini	2011	Perception	16	fMRI	Emotion non-specific	N	8	8	16	N	21.75	–	–	–	–	–	–	–
89	Pfordresher	2014	Perception	20	fRMI	Melody	Y	4	16	20	Y	24	–	13.4	–	–	–	–	–
90	Ragert	2014	Perception	17	fMRI	Melody	N	9	8	17	Y	26.12	4.2	16.44	5.92	–	–	–	–
91	Reiterer	2008	Perception	17	fMRI	Timbre	N	8	9	17	N	25.2	–	–	–	–	–	–	–
92	Rogalsky	2011	Imagery	20	fMRI	Melody	Y	9	11	20	N	22.6	–	3.5	–	–	–	–	–
93	Sammler	2010	Imagery	12	fMRI	Familiarity	Y	6	6	12	N	29	–	–	–	–	–	–	–
94	Schmithorst	2005	Perception	15	fMRI	Melody	Y	4	11	15	N	37.8	15.2	–	–	–	–	–	–
95	Schön	2010	Perception	10	fMRI	Non-specific	N	2	8	10	N	24	–	3	–	–	–	–	–
96	Schwenzer	2011	Perception	16	fMRI	Pitch	N	8	8	16	N	–	–	–	–	–	–	–	–
97	Shany	2019	Perception	40	fMRI	Emotion +	N	18	22	40	B	25.5	3.6	6	–	–	–	–	–
98	Sikka	2015	Perception	40	fMRI	Familiarity	N	0	40	40	N	–	–	–	–	–	–	–	–
99	Singer	2016	Perception	40	fMRI	Emotion non-specific	N	18	22	40	B	25.5	3.6	12.3	4.7	–	–	–	–
100	Skouras	2014	Production	32	fMRI	Emotion non-specific	Y	14	18	32	N	22.93	2.75	3.62	5.04	–	–	–	–
101	Spada	2014	Production	20	fMRI	Melody + harmony	N	6	14	20	Y	30.29	7.02	21	6.86	12	–	–	–
102	Tabei	2015	Perception	17	fMRI	Emotion non-specific	N	7	10	17	N	21.4	2	–	–	–	–	–	–
103	Tachibana	2010	Imagery + Production	12	fMRI	Pitch	N	7	5	12	N	24	–	9.1	5.4	–	–	–	–
104	Taruffi	2017	Perception	24	fMRI	Emotion + and −	N	12	12	24	N	25.3	–	–	–	–	–	–	–
105	Tervaniemi	2000	Perception	30	PET	Harmony	N	30	0	30	N	26.2	–	–	–	–	–	–	–
106	Tervaniemi	2006	Perception	17	fMRI	Timbre	N	9	8	17	N	25.1	–	–	–	–	–	–	–
107	Thaut	2008	Perception	12	fMRI	Rhythm + beat + metre	N	9	3	12	Y	26.1	1.8	–	–	–	–	7	–
108	Tillmann	2003	Perception	15	fMRI	Tonality	Y	6	9	15	Y	19.4	–	8.2	2.54	–	–	–	–
109	Tillmann	2006	Perception	20	fMRI	Timbre/familiarity	Y	9	11	20	Y	25.1	–	4.5	5.3	–	–	–	–
110	Toiviainen	2014	Perception	15	fMRI	Various	N	10	5	15	N	25.7	5.2	–	–	–	–	–	–
111	Trost	2012	Perception	15	fMRI	Emotion + and −	N	8	7	15	N	28.8	9.9	–	–	–	–	–	–
112	Trost	2014	Production	18	fMRI	Harmony	Y	7	11	18	N	25.8	7.5	5	–	–	–	-	-
113	Tsai	2010	Perception	12	fMRI	Rhythm + beat + metre	Y	2	10	12	N	–	–	–	–	–	–	–	–
114	Tsai	2012	Perception	15	fMRI	Rhythm + timbre	N	12	3	15	Y	22	–	6.4	3.1	–	–	–	–
115	Tsai	2018	Perception	14	fMRI	Pitch	N	3	11	14	N	–	–	–	–	–	–	–	–
116	Tsai	2019	Production	16	fMRI	Pitch/harmony	Y	5	11	16	N	22.44	1.58	–	–	–	–	–	–
117	Uhlig	2013	Perception	17	fMRI	Melody	Y	9	8	17	Y	26.12	4.2	16.44	5.92	–	–	–	–
118	Villarreal	2013	Perception	24	fMRI	Rhythm + creativity	N	9	15	24	N	21.9	1.6	–	–	–	–	–	–
119	Vuust	2011	Perception	18	fMRI	Rhythm	N	14	4	18	Y	29	1	–	–	8	–	7	–
120	Wallmark	2018	Perception	15	fMRI	Emotion	N	7	8	15	N	19.1	0.72	–	–	–	–	–	–

Continued

	Author	Year	Domain	n	Method	Feature	Familiarity	Males	Females	Total	Musicians	Age		Education (years)		Onset (years)		Training (hr/week)	
												Mean	± SD	Mean	± SD	Mean	± SD	Mean	± SD
121	Watanabe	2008	Perception	18	fMRI	Memory	N	12	6	18	N	22.4	0.4	–	–	–	–	–	–
122	Whitehead	2018	Perception	24	fMRI	Melody	N	13	11	24	N	25.5	–	4.2	4.73	–	–	–	–
123	Wilson	2010	Perception	26	fMRI	Melody	Y	11	15	26	B	31.7	12.4	–	–	–	–	–	–
124	Yoo	2001	Perception	12	fMRI	Pitch	N	8	4	12	N	29.9	5.8	–	–	–	–	–	–
125	Zarate	2008	Perception	24	fMRI	Pitch	Y	12	12	24	B	–	–	10	3.39	–	–	–	–
126	Zarate	2010	Perception	9	fMRI	Pitch	Y	3	6	9	Y	23	3.93	11	4.28	–	–	–	–
127	Zatorre	1994	Perception	12	PET/fMRI	Melody	N	6	6	12	N	–	–	–	–	–	–	–	–
128	Zatorre	1996	Perception	12	PET	Melody	N	6	6	12	N	22	–	–	–	–	–	–	–
129	Zatorre	2010	Perception	12	fMRI	Melody	Y	5	7	12	Y	23	–	16	–	–	–	–	–
130	Zvyagintsev	2013	Perception	15	fMRI	Melody	N	8	7	15	N	25.1	5.7	–	–	–	–	–	–

Table 1. Characteristics of the studies included in meta-analysis. *Y* yes, *N* no, *B* both.

Primary outcomes: ALE meta-analyses of music perception, production, and imagery. *Music perception.* The music perception ALE meta-analysis included 1898 foci, 105 experiments and 2035 subjects. Significant peak clusters resulted in the following areas: (1) right superior temporal gyrus (x = 52, y = − 20, z = 4) extending over insula and inferior frontal gyrus; (2) left superior frontal gyrus (x = − 54, y = − 16, z = 2) extending over transverse temporal gyrus, the insula and precentral gyrus; (3) left medial frontal gyrus (x = − 2, y = − 2, z = 66) including the superior frontal gyrus and the right medial frontal gyrus; (4) right lentiform nucleus (putamen) (x = 22, y = 8, z = 6) and the caudate body; (5) left lentiform nucleus (putamen) (x = − 22, y = 4, z = 6) and the caudate body; (6) left cerebellum (lobule III) (x = − 28, y = − 64, z = − 26); (7) left insula (x = − 32, y = 18, z = 10) including the inferior frontal gyrus and (8) right precentral gyrus (x = 54, y = 0, z = 46). Such results show consistency of cortical and subcortical areas presumably organized in a hierarchical manner processing both bottom-up and top-down musical information (Fig. 1, Table 2).

Music production. The music production ALE meta-analysis included 499 foci, 19 experiments and 292 subjects. Significant peak clusters resulted in the following areas: (1) right superior temporal gyrus (x = 66, y = 24, z = 10) extending over Wernicke's area, anterior and posterior transverse temporal area; (2) left precentral gyrus (x = − 48, y = 0, z = 42) extending over precentral gyrus, inferior parietal lobule and post central gyrus; (3) left medial frontal gyrus (x = − 10, y = − 10, z = 52) including cingulate gyrus; (4) left superior temporal gyrus (x = − 42, y = 28, z = 6); (5) right precentral gyrus (x = 54, y = 4, z = 32) extending over the inferior frontal gyrus. These results evidence not only the obvious involvement of the somatomotor system, but also limbic and executive cognitive functions while playing music. Interestingly, similar areas were found in music perception, further evidencing the integration of incoming information with high-order functions and motor behaviour (Fig. 1, Table 2).

Music imagery. The music imagery ALE meta-analysis included 263 foci, 15 experiments and 189 subjects. Significant peak clusters resulted in the following areas: (1) left medial frontal gyrus (x = 0, y = 6, z = 58); (2) left superior parietal lobule (x = − 34, y = − 58, z = 56) extending over angular gyrus and inferior parietal lobule; (3) left thalamus (x = − 14, y = − 14, z = 8); (5) left precentral gyrus (x = 52, y = 10, z = 4) encompassing the middle frontal gyrus. These results show that music imagination recruits motor areas involved in the generation of movements such as the premotor cortex and supplementary motor area; areas of the basal ganglia which are also involved in facilitation of movements; and parietal areas involved in perceptual-motor coordination and theory-of-mind (Fig. 1, Table 2).

Contrast analyses. A subsequent contrast ALE meta-analysis was conducted to compare the primary outcomes, by means of conjunction and subtraction. The purpose of this analysis was to identify common and distinct brain activity between perception, production and imagery modalities of music cognition (Supplementary Table 3).

Music perception vs music production. Clusters identified as likely to be active during both music perception and production (conjunction) include the right superior temporal gyrus (BA42), bilateral superior temporal gyrus (BA13), right superior temporal gyrus (BA22), left superior temporal gyrus (BA41), insula, right transverse temporal gyrus (BA42), left medial frontal gyrus (BA6) and inferior frontal gyrus (BA9). The contrast analysis (subtraction) yielded areas specific to perception (perception–production) and production (production–perception). Areas more likely to be active in music perception are the superior temporal gyrus (BA21) and superior temporal gyrus (BA22). Conversely, areas more likely to be active in music production include bilateral post central gyrus (BA2, BA40), bilateral precentral gyrus (BA4, BA6), bilateral superior, middle and medial frontal gyrus (BA6), left cingulate gyrus (BA24, BA32), left superior temporal gyrus (BA41, BA13, 22), left inferior parietal lobule (BA40), left insula (BA13), cerebellum (culmen) and the lentiform nucleus.

Music production vs music imagery. Clusters that are presumably active during both music production and music imagery (conjunction) includes the left medial frontal gyrus (BA6). The contrast analysis (subtraction) resulted in areas either specific to music production (production-imagery) or to music imagery (imagery-pro-

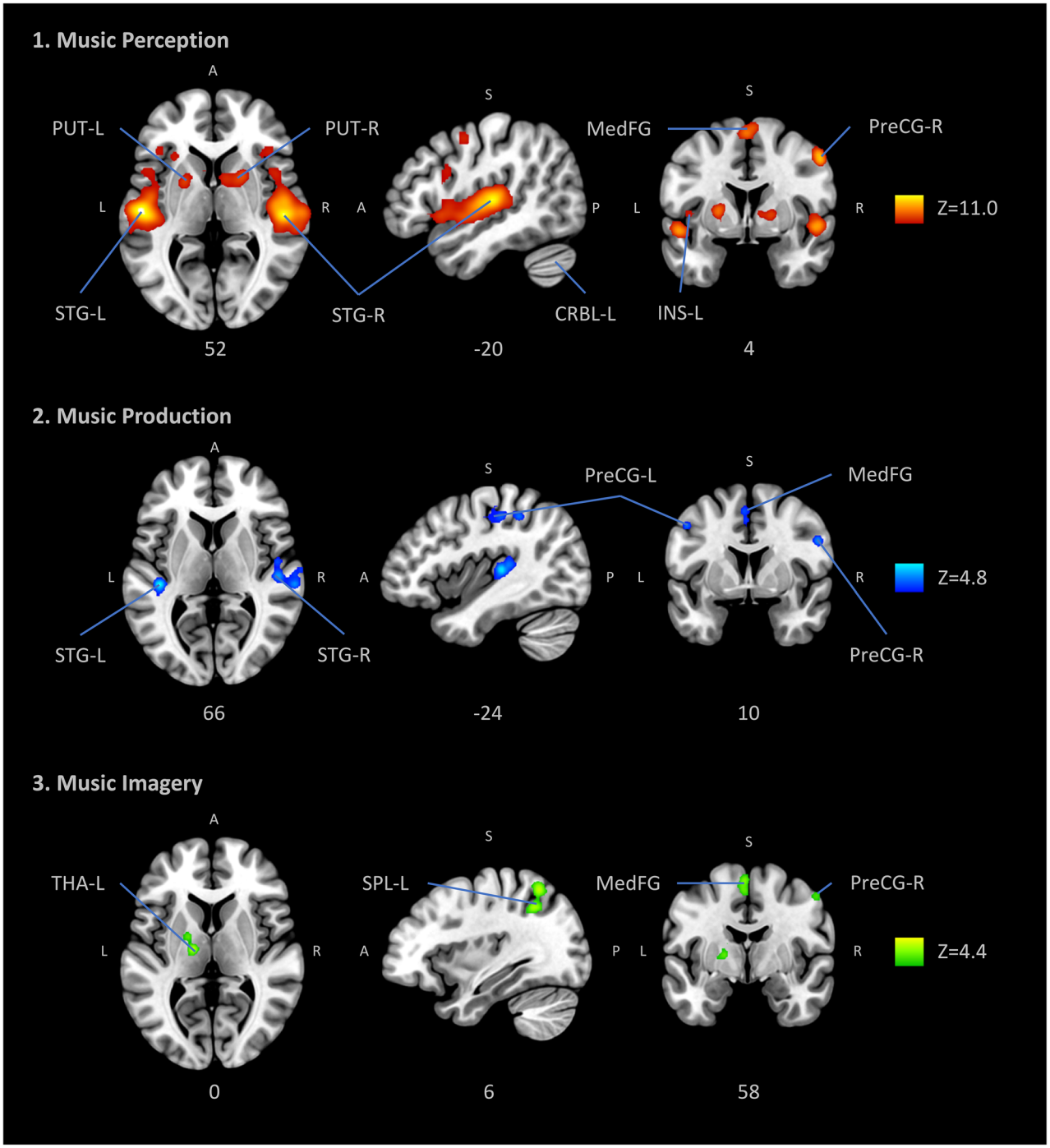

Figure 1. Anatomic likelihood estimation meta-analytic results for studies of music perception, production and imagery, at cluster level inference $p < 0.05$ (FWE). The primary outcomes included ALE meta-analyses of music perception, music production, and music imagery, independently. ROIs: *CRBL* cerebellum, *INS* insula, *MedFG* medial frontal gyrus, *PreCG* precentral gyrus (primary motor cortex or M1), *PUT* putamen, *SPL* superior parietal lobule, *STG* superior temporal gyrus (primary auditory cortex), *THA* thalamus, *L* left, *R* right, *Z* peak Z-value. Figure created with Mango (http://rii.uthscsa.edu/mango//userguide.html).

duction). Areas more likely to be active in music production include the left superior temporal gyrus (BA13, BA41, BA22), left transverse temporal gyrus (BA41), left postcentral gyrus (BA40) and the culmen. Areas more likely to be active while imagining music include the left superior parietal lobule (BA7), angular gyrus (BA39), medial frontal gyrus (BA6) and the precentral gyrus (BA6).

Music imagery vs music perception. Clusters likely active during both music imagery and music perception (conjunction) include the left medial frontal gyrus (BA6), right precentral gyrus (BA6) and the globus pallidus.

Cluster number	Volume (mm^3)	MNI coordinates			ALE	P	Z	Label (Side region BA)
		x	y	z				
1. Music perception: 1898 foci, 105 experiments, 2035 subjects								
1	24,296	52	−20	4	1E−01	7E−27	10.7	R Superior Temporal Gyrus BA13
		54	−4	−4	8E−02	3E−17	8.4	R Superior Temporal Gyrus BA22
		64	−30	10	5E−02	2E−10	6.2	R Superior Temporal Gyrus BA42
		48	14	−4	4E−02	2E−07	5.1	R Insula BA13
		42	12	6	3E−02	8E−06	4.3	R Insula BA13
		42	26	2	3E−02	1E−05	4.3	R Inferior Frontal Gyrus BA13
		48	10	26	3E−02	8E−05	3.8	R Inferior Frontal Gyrus BA9
2	23,864	−54	−16	2	1E−01	1E−28	11.0	L Superior Temporal Gyrus BA22
		−40	−30	12	7E−02	4E−14	7.5	L Transverse Temporal Gyrus BA41
		−48	−34	12	6E−02	7E−12	6.8	L Superior Temporal Gyrus BA41
		−50	14	−2	4E−02	3E−07	5.0	L Insula BA13
		−54	6	12	3E−02	7E−05	3.8	L Precentral Gyrus BA6
3	6136	−2	−2	66	5E−02	7E−09	5.7	L Medial Frontal Gyrus BA6
		−2	12	54	5E−02	7E−09	5.7	L Superior Frontal Gyrus BA6
		4	−2	68	5E−02	2E−08	5.5	R Medial Frontal Gyrus BA6
		2	16	48	5E−02	2E−08	5.5	L Medial Frontal Gyrus BA32
4	3328	22	8	6	6E−02	6E−11	6.4	R Putamen
		10	6	4	4E−02	8E−07	4.8	R Caudate Body
5	2304	−22	4	6	5E−02	1E−09	6.0	L Putamen
		−8	10	0	4E−02	7E−07	4.8	L Caudate Head
6	1920	−28	−64	−26	5E−02	1E−08	5.6	L Cerebellum
7	1816	−32	18	10	4E−02	3E−07	5.0	L Insula BA13
		−30	24	0	3E−02	6E−06	4.4	L Insula BA13
		−42	26	4	3E−02	4E−05	3.9	L Inferior Frontal Gyrus BA13
8	1728	54	0	46	6E−02	4E−12	6.8	R Precentral Gyrus BA4
2. Music production: 499 foci, 19 experiments, 292 subjects								
1	3512	66	−24	10	2E−02	7E−07	4.8	R Superior Temporal Gyrus BA42
		50	−20	6	2E−02	7E−07	4.8	R Superior Temporal Gyrus BA13
		64	−26	4	2E−02	7E−07	4.8	R Superior Temporal Gyrus BA22
		62	−22	2	2E−02	4E−06	4.5	R Superior Temporal Gyrus BA41
		50	−10	4	2E−02	2E−05	4.1	R Insula BA13
		64	−28	14	2E−02	2E−05	4.1	R Superior Temporal Gyrus BA42
		68	−16	8	2E−02	4E−05	3.9	R Transverse Temporal Gyrus BA42
2	2920	−48	0	42	2E−02	8E−07	4.8	L Precentral Gyrus BA6
		−40	−40	46	2E−02	5E−06	4.4	L Inferior Parietal Lobule BA40
		−36	−46	48	2E−02	7E−06	4.3	L Inferior Parietal Lobule BA40
		−40	−26	46	2E−02	1E−05	4.2	L Inferior Parietal Lobule BA40
		−44	−22	50	2E−02	5E−05	3.9	L Postcentral Gyrus BA2
		−44	−14	44	2E−02	1E−04	3.6	L Precentral Gyrus BA4
3	2080	−10	−10	52	2E−02	1E−06	4.7	L Medial Frontal Gyrus BA6
		0	−4	58	2E−02	9E−06	4.3	L Medial Frontal Gyrus BA6
		−4	4	54	2E−02	2E−05	4.1	L Medial Frontal Gyrus BA6
		−2	2	46	2E−02	5E−05	3.9	L Cingulate Gyrus BA24
4	1912	−42	−28	6	3E−02	3E−09	5.8	L Superior Temporal Gyrus BA41
5	1072	54	4	32	2E−02	3E−07	5.0	R Precentral Gyrus BA6
		52	10	24	2E−02	8E−05	3.8	R Inferior Frontal Gyrus BA9
3. Music imagery: 263 foci, 15 experiments, 189 subjects								
1	3256	0	6	58	2E−02	6E−09	5.7	L Medial Frontal Gyrus BA6
		−2	0	62	2E−02	1E−07	5.2	L Medial Frontal Gyrus BA6
2	3128	−34	−58	56	2E−02	1E−07	5.1	L Superior Parietal Lobule BA7
		−32	−54	42	2E−02	2E−07	5.1	L Angular Gyrus BA39
		−40	−64	52	2E−02	8E−06	4.3	L Inferior Parietal Lobule BA7
3	1384	−14	−14	8	2E−02	2E−07	5.1	L Thalamus
		−18	−4	6	2E−02	2E−05	4.2	L Thalamus

Continued

Cluster number	Volume (mm^3)	MNI coordinates			ALE	P	Z	Label (Side region BA)
		x	y	z				
4	944	54	0	50	2E−02	1E−05	4.2	L Precentral Gyrus BA6
		52	10	44	1E−02	7E−05	3.8	L Middle Frontal Gyrus BA6

Table 2. Anatomic likelihood estimation meta-analytic results for music perception, production and imagery at cluster level inference $p < 0.05$ (FWE). *ALE* anatomic likelihood estimation, *BA* Brodmann area, *P* p-value, *Z* peak z-value, *R* right, *L* left.

The contrast analysis (subtraction) resulted in areas either specific to music imagery (imagery–perception) or to music perception (perception–imagery). Areas more related to music imagery include left inferior parietal lobule (BA7, BA40), right middle and left medial frontal gyrus (BA6), right precentral gyrus (BA4) and the ventral lateral and ventral posterior lateral thalamus. Finally, areas more likely to be active during music perception include the bilateral superior temporal gyrus (BA22), left transverse temporal gyrus (BA41), right claustrum, left insula (BA13) and left caudate head.

Meta-analytic connectivity modelling (MACM). MACM was performed to functionally segregate the behavioural contribution and the patterns of co-activation of each music-related region-of-interest (ROI) resulted from the primary outcomes (n = 17). The ROIs were imported into the BrainMap database separately, to identify studies reporting activation within each ROI boundary (Supplementary Table 4). The functional characterization of each ROI is detailed in Supplementary Table 5 and include the behavioural domains of action, perception, emotion, cognition and interoception.

Music perception MACM. The right superior temporal gyrus ROI (1a) showed co-activation with left superior temporal gyrus, right claustrum, left medial frontal gyrus, right precentral gyrus, left insula, and left cerebellum. Relevant behavioural domains within its boundaries include execution, language, music, and auditory perception; and experimental paradigms including emotion induction, finger tapping, music comprehension, music production, and oddball, phonological, pitch, semantic, and tone discrimination.

The left superior temporal gyrus ROI (1b) showed co-activation with right superior temporal gyrus and right insula. Relevant behavioural domains within its boundaries include execution, language, music, positive emotion, and auditory perception; and experimental paradigms including emotion induction, finger tapping, music comprehension, music production, and oddball, phonological, pitch, semantic, and tone discrimination.

The left medial frontal gyrus ROI (1c) showed co-activation with left precentral gyrus, bilateral thalamus, bilateral cerebellum, bilateral superior temporal gyrus, and left superior parietal lobule. Relevant behavioural domains within its boundaries include execution, attention, music, and auditory perception; and experimental paradigms including emotion induction, finger tapping, music comprehension, music production, and pitch, semantic, tactile, and tone discrimination.

The right putamen ROI (1d) showed co-activation with left putamen, left medial frontal gyrus, and right inferior parietal lobule. Relevant behavioural domains within its boundaries include execution, attention, language, working memory, music, reasoning, reward, and auditory perception; and experimental paradigms including encoding, finger tapping, music comprehension, music production, reward, and pitch, semantic, tactile, and tone discrimination.

The left putamen ROI (1e) showed co-activation with right putamen, left medial frontal gyrus, bilateral superior temporal gyrus, and left cerebellum. Relevant behavioural domains within its boundaries include execution, attention, language, working memory, music, reward, and auditory perception; and experimental paradigms including emotion induction, finger tapping, music comprehension, and phonological, pitch, semantic, and tone discrimination.

The left cerebellum ROI (1f) showed co-activation with right cerebellum, bilateral insula, left medial frontal gyrus, bilateral superior temporal gyrus, left precentral gyrus, and right thalamus. Relevant behavioural domains within its boundaries include execution, attention, language, music, and auditory perception; and experimental paradigms including emotion induction, music comprehension, music production, and pitch, semantic, tactile, and tone discrimination.

The left insula ROI (1g) showed co-activation with right insula, left medial frontal gyrus, right inferior frontal gyrus, left inferior parietal lobule, left precentral gyrus, and right middle frontal gyrus. Relevant behavioural domains within its boundaries include attention, language, working memory, music, reasoning, reward, and auditory perception; and experimental paradigms including emotion induction, finger tapping, music comprehension, reward, and oddball, phonological, semantic, and tone discrimination.

The right precentral gyrus ROI (1h) showed co-activation with left precentral gyrus, left medial frontal gyrus, bilateral superior temporal gyrus, bilateral putamen, and right thalamus. Relevant behavioural domains within its boundaries include execution, attention, language, music, and auditory perception; and experimental paradigms including emotion induction, finger tapping, music comprehension, music production, reward, and pitch, semantic, and tone discrimination.

Music production MACM. The right superior temporal gyrus ROI (2a) showed co-activation with left superior temporal gyrus, left medial frontal gyrus, left insula, left thalamus, and right precentral gyrus. Relevant behavioural domains within its boundaries include execution, attention, language, music, and auditory perception; and experimental paradigms including emotion induction, finger tapping, music comprehension, music production, reward, and oddball, phonological, pitch, semantic, and tone discrimination.

The left precentral gyrus ROI (2b) showed co-activation with right precentral gyrus, left medial frontal gyrus, left inferior parietal lobule, left middle temporal gyrus, left fusiform gyrus, right superior parietal lobule, and right cerebellum. Relevant behavioural domains within its boundaries include execution, attention, language, working memory, music, reasoning, and visual motor perception; and experimental paradigms including finger tapping, music comprehension, music production, and phonological, pitch, semantic, and tactile discrimination.

The left medial frontal gyrus ROI (2c) showed co-activation with left precentral gyrus, right cerebellum, left postcentral gyrus, and left claustrum. Relevant behavioural domains within its boundaries include execution, attention, and reward; and experimental paradigms including finger tapping, music comprehension, music production, reward, semantic discrimination, and tactile discrimination.

The left superior temporal gyrus ROI (2d) showed co-activation with right superior temporal gyrus, and left putamen. Relevant behavioural domains within its boundaries include execution, attention, language, music, social cognition, and auditory perception; and experimental paradigms including emotion induction, finger tapping, music comprehension, music production, and oddball, phonological, pitch, semantic, and tone discrimination.

The right precentral gyrus ROI (2e) showed co-activation with left precentral gyrus, anterior cingulate cortex, left cerebellum, bilateral superior parietal lobule, bilateral inferior parietal lobule, right insula, and right thalamus. Relevant behavioural domains within its boundaries include execution, language, reasoning, music, and social cognition; and experimental paradigms including emotion induction, finger tapping, music comprehension, music production, phonological discrimination, and pitch discrimination.

Music imagery MACM. The left medial frontal gyrus ROI (3a) showed co-activation with bilateral insula, left superior parietal lobule, right inferior parietal lobule, and bilateral cerebellum. Relevant behavioural domains within its boundaries include execution, attention, language, memory, music, reasoning, and auditory perception; and experimental paradigms including emotion induction, finger tapping, music comprehension, music production, reward, and phonological, pitch, semantic, syntactic, tactile, and tone discrimination.

The superior parietal lobule ROI (3b) showed co-activation with right superior parietal lobule, left anterior cingulate cortex, right inferior frontal gyrus, and bilateral insula. Relevant behavioural domains within its boundaries include execution, attention, language, working memory, music, reasoning, reward, and auditory perception; and experimental paradigms including emotion induction, finger tapping, music comprehension, reward, and pitch, semantic, and tactile discrimination.

The left thalamus ROI (3c) showed co-activation with right thalamus, anterior cingulate cortex, left precentral gyrus, left inferior frontal gyrus, right insula, and right cerebellum. Relevant behavioural domains within its boundaries include execution, attention, working memory, music, reward, auditory perception; and experimental paradigms including emotion induction, finger tapping, motor learning, music comprehension, music production, reward, and oddball, phonological, tactile, and tone discrimination.

d. The right precentral gyrus ROI (3d) showed co-activation with left precentral gyrus, left medial frontal gyrus, right insula, bilateral putamen, and bilateral superior temporal gyrus. Relevant behavioural domains within its boundaries include execution, music, negative emotion, and auditory perception; and experimental paradigms including emotion induction, finger tapping, music comprehension, music production, and phonological, pitch, semantic, tactile, and tone discrimination.

Discussion

In the present study, we have conducted ALE meta-analyses of neuroimaging studies with a total of 2660 foci, 139 experiments, and 2516 subjects, to realise a comprehensive picture of the top-down and bottom-up mechanisms required by music processing. Our main results, whose robustness is guaranteed by the large numbers of studies included in the analyses, provide a complex and appetising picture of the brain activity underlying music processing. We show that music perception and music production rely on similar brain activation involving auditory cortices and sensorimotor cortices. In turn, music imagery presents a particular and unique activation of parietal and motor regions. Finally, our primary outcomes were complemented with contrast analyses and meta-analytic connectivity modelling, describing in full the complexity of the brain areas underlying music processing.

Characteristics of included studies. The publications included in this comprehensive systematic review and meta-analysis reported a clear research question, inclusion and exclusion criteria for participants, description of methods and explicit results. In line with reproducibility efforts, the included studies used state-of-the-art techniques and computational MRI tools important for the support of standardization of neuroimaging studies. However, some works lacked important demographic data such as the years of education, age of musical training onset, and current time of musical practice, which may have an effect on behavioural tasks and neuroimaging data.

Music perception. *Temporal lobe areas.* As expected, the main result concerning music perception involved temporal regions and especially auditory cortices. Indeed, we found convergence in various temporal lobe areas assumed to be related to auditory feature processing as well as semantic and conceptual aspects of music. Namely, primary auditory regions of Heschl's gyrus (BA41, BA42), as well as secondary auditory area,

namely Wernicke's area (BA22). Mirrored by previous studies, these areas have been seen crucial for processing of auditory information including pattern recognition[25] and syntax processing[26]. Hence, activation of both primary and secondary areas may reflect a hierarchical computation of auditory processing including lower-level extraction followed by alignment to preconstructed predictions in association areas. Our MACM results show that the superior temporal gyrus where the primary auditory cortex lies, co-activates with motor and pre-motor areas, insula, and cerebellum, supporting the idea that a large network is established while listening to music involving audio-motor coupling and limbic and paralimbic processing of emotional content.

Audio-motor coupling. We also found convergence in motor-associated areas, assumingly related to temporal processing, inherent audio-motor coupling and entrainment. These include areas within the basal ganglia (caudate, putamen, globus pallidus and thalamus), primary motor areas (precentral gyrus BA4), supplementary motor areas (premotor cortex, supplementary motor area BA6), and the cerebellum (lobule III). As music unfolds over time and does so at a specific periodicity (rhythm, beat, metre), it is not surprising to see brain activation in areas related to temporal processing in the motor domain[27]. Furthermore, previous studies have proposed significant connectivity within the fronto-temporal-cerebellar network during the perceptual analysis of music listening[28,29]. In concordance with neuroimaging results, electroencephalography (EEG) studies have found synchronised neural activity in the gamma-band within this network, in response to rhythmic expectancy[30,31]. Additionally, specific neuronal firing patterns were found in those areas with respect to time comprehension[32]. These include ramping spike activity, which refers to a consistent increase or decrease in neuronal firing rate, spectral responses and network-produced ensemble patterns[33]. In relation to the lentiform nucleus, the putamen has been proposed to be important for internal beat perception[34] and for generation of internalised groupings or chunking of action representations[35]. MACM revealed an extensive network that can represent audio-motor coupling with areas such as primary auditory and motor cortices, premotor cortex, basal ganglia, insula, and cerebellum.

Broad activation of the insular cortex. Further, our ALE analysis outlined an arguably broad activation of the insular cortex (BA13) extending under the superior frontal gyrus, sub-lobar insular and superior temporal gyrus. The insula cortex, folded deep within the lateral sulcus and covered by the frontoparietal and temporal operculum[36], has been primarily divided into two parts: the anterior and posterior insula[37]. However, more than a dozen subdivisions have been identified[38]. With rich connectivity with both cortical and subcortical structures, the insular cortex is seen as an anatomical integration hub with rich connectivity and heavy cross-modal input[39]. Within our MACM analysis, this is particularly evident with the functionally connectivity of the insula spreading across both cortical and subcortical structures, likely reflecting the breadth of its function as a multimodal integration center. As categorised by Kurth and colleagues[40] in a meta-analysis of close to 1800 experiments, the insular cortex serves a number of categorically distinct functions including sensorimotor, olfacto-gustatory, socio-emotional and cognitive functions. With respect to music perception, such rich functional diversity aligns with previous findings presenting the role of the insular cortex in auditory processing including allocation of auditory attention[41], time perception and interceptive salience[42] as well as music evoked emotion and changes in physiological arousal[43]. With these in mind, a broad activation of the insular cortex in our results appraises the complexity and multiplicity of music as a stimulus to the human ear.

Music production. *Motor areas.* Playing an instrument or engaging in music production requires extremely intricate motor control functions as well as the ability to internalise and portray emotional meaning. Accordingly, and unsurprisingly, with regards to our music production analysis, we found convergence in areas related to motor planning and execution. Namely, areas within the precentral gyrus, (BA4), middle frontal gyrus (BA6) and the supramarginal gyrus (BA40). Aside from their obvious motor function in relation to instrument playing, these areas are also likely to be involved in various sensory-motor coupling. Previous studies have repeatedly outlined the PMC and the planum temporale to be the crucial players in various sensorimotor transformations[44]. Activation within the somatosensory association area such as the supramarginal gyrus has been previously linked to tactile interpretation and limb location[45] crucial for use of an instrument. In relation to music, the rostral dPMC has been recruited during increasingly more complex rhythms[34] and saliency[46]. This means that the PMC could be responsible for modifying motor sequences to reflect auditory sequential patterns in a predictive and anticipatory manner.

Limbic areas. Aside from motor dexterity crucial for the technical component of instrument playing, the performance must engage the listener with loaded emotional content. Although with reduced breadth of activation compared to motor areas, our results point to a number of regions involved in the processing and conveyance of emotional material including the anterior cingulate gyrus (BA24) and insula (BA13). Previous findings of neuroanatomical studies have outlined the temporal lobe having strong connectivity with the anterior cingulate gyrus[47]. Further, for an appropriate emotional response the anterior cingulate has shown to require specific sensory information[44]. Thus, the anterior cingulate gyrus may be crucial for top-down control of emotional meaning and bottom-up processing of emotional content. In complement, numerous neuroimaging studies have outlined the activation of cingulate gyrus during musical performance[48,49]. Interestingly, an emerging body of research is suggesting a role of the cerebellum in cognition and emotion with reciprocal connections to various limbic system regions. The fastigial nucleus in particular has been shown to have extensive topographic connections with non-motor systems including the limbic system encompassing areas such as the anterior cingulate gyrus[50].

Music imagery. *Frontal regions.* *While imagining music*, our results point to a convergence in motor-associated areas within the medial frontal gyrus (supplementary motor area or SMA), precentral gyrus, putamen and thalamus. As previously mentioned, neuroimaging studies on rhythm, metre and beat have outlined the medial frontal gyrus, superior frontal gyrus along with the basal ganglia in the context of temporal feature processing and inherent audio-motor coupling[51]. In relation to imagery, activation within the SMA is particularly interesting, being repeatedly found in various imagery domains including visual imagery[52]. Furthermore, this region is also found in relation to imagination in hallucination studies[53]. Within our results, the SMA was specific to imagery suggesting its distinct role in mental manipulation and generation of internalised representations also during music.

Parietal regions. The main result from our ALE meta-analysis show music imagery to recruit parietal regions including the superior and inferior parietal lobule as well as the angular gyrus. With regards to the parietal lobules (BA7), this result is far from surprising. These areas have been previously discussed in the context of auditory attention[54] and working memory[55]. Specific to imagery, previous research on motor imagery has reported the parietal lobule as a key region in several documentations using rTMS[56] and fMRI[57]. Furthermore, activation within the angular gyrus (BA39) compliments the aforementioned parietal activation. Previous findings have suggested its role in memory retrieval[58], working memory[59], and perspective taking[60]. Further, our MACAM analysis specific to the parietal areas have revealed a distinct connection pattern with a sequence of areas including frontoparietal and motor regions. Together, their joint role could be interpreted as key in initiating and formulating the experience of imagery involving appropriate cognitive abilities.

Importantly, what must be noted is that despite several authors describing imagery to recruit secondary auditory areas, we found no temporal area activation as it did not survive the correction for multiple comparisons. However, such temporal areas were only found in the uncorrected analysis together with other limbic and paralimbic areas. This may be due to variability of designs in the imagery studies, individual differences in imagery vivacity or a low number of studies.

Comparison between music perception, production, and imagery. A direct comparison of music perception, production and imagery for each of the respective pairs revealed several indications of music processing specificity likely resulting from a bidirectional organisation of the brain. Music imagery showed a unique activation of left superior and inferior parietal lobule (BA7). Previously, studies outlined these areas to be key for top-down or goal directed attentional orienting[61], and specific to imagery, the superior parietal lobule has been named a potential candidate for top-down generation and maintenance of domain-unspecific mental images[62]. Crucially, both music perception and production, when contrasted with music imagery, showed similar broad activation of temporal and frontal areas associated with auditory processing and motor planning and execution. This is particularly interesting as to show that both perception and production must involve both top-down and bottom-up processes, thus, supporting the embodied and predictive approach which appreciates the human motor system and its actions as a reciprocal part of perception and cognition itself[63].

Limitations and future perspectives. Although our meta-analyses comprise a large number of studies and thus arguably provide reliable results, some limitations should be mentioned. First, by definition meta-analyses rely on the combination of the data from previous studies and therefore are highly dependent on the quality of such studies. In our case, we utilised studies that possessed high-quality standards, making feasible data extraction and conduction of statistical analysis. However, we should highlight that a thorough description of the methods and an even stronger focus on the description of the experimental settings would be highly beneficial for the field and thus we encourage future studies to further increase their already high-level quality. Second, we relied on the assumptions of the GingerALE method that was used in our study. Therefore, the accuracy of our findings necessarily relies on the statistical estimation of coordinate-based anatomic foci (input) which were treated as spatial probability distributions centred at the given coordinates. As conceivable, we could not control the heterogeneity of the methods used in the studies included in the meta-analysis (ranging from pre-processing software, statistical thresholds, smoothing, and participants' characteristics) and we acknowledge that this fact could represent a potential confounder for our analyses.

As in traditional meta-analyses, we have assessed the publication bias[64] and revealed that our results are robust as the clusters obtained from the primary outcomes presented FSN values within lower and upper boundaries. Importantly, this outcome proves a robust convergence of the foci emerged from our analyses even when the total amount of studies was small, particularly in the music production and imagery CBMA's (details in Supplementary Table 6).

With regard to future perspectives, we believe that the neuroscientific community would benefit from designing new ad-hoc studies to directly investigate the relationship between brain activity and different features of music processing such as music perception, production and imagination. Although our results provide a robust picture of the brain networks recruited by those different processes, carefully designed experiments which systematically modulate and compare stimuli and conditions would further refine the knowledge on this broad and complex topic.

Conclusions

In this comprehensive review, we have conducted a series of neuroimaging meta-analyses on decades of studies investigating the brain activity underlying music processing. We have shed new light on how the brain processes and integrates top-down and bottom-up mechanisms during complex multi-modal human activities such as music. The outcomes of our synthesis highlight three main concurrent music procedures, known as music

perception (both top-down and bottom-up), music production (both top-down and bottom-up) and imagery (top-down). Our results show that the brain relies on different structures and mechanisms to process similar musical information. Indeed, music perception and music production depend on auditory cortices, sensorimotor cortices and cerebellum. Differently, music imagery shows a key recruitment of parietal regions. Taken together, our findings provide robust evidence that the brain requires different structures to process similar information which is made available either by the interaction with the environment (i.e., bottom-up) or by internally generated content (i.e., top-down).

Methods and materials

Search strategy, screening, and extraction. This comprehensive systematic review and meta-analysis followed procedures from the Cochrane Handbook for Systematic Reviews[65], and from the Center for Reviews and Dissemination (https://www.york.ac.uk/crd/). The review protocol was pre-registered in PROSPERO (CRD42019140177), and was carried out in accordance with the PRISMA statement[66]. A systematic literature search was conducted in PubMed, Scopus and PsycInfo for articles published up to February 26th, 2021. The search strategy for music perception, production and imagery was developed using keywords and MeSH terms (details in Supplementary Information p. 2).

To ensure concurrence between the research question and data extracted, the following inclusion criteria was established: (1) healthy adults (18–80 age range); (2) studies investigating music listening, music playing, or music imagery, with no restrictions on music genre, parameter, or instrument; (3) using fMRI or PET; (4) whole-brain analyses; and (5) results reported in stereotactic coordinates either Talairach or Montreal Neurological Institute (MNI) three-dimensional-coordinate system. Studies were excluded using the following criteria: (1) review articles with no original experimental data, (2) populations with any form of disorder or disability; (3) paediatric population; (4) auditory stimuli not related to music; (5) functional connectivity analyses; (6) region-of-interest (ROI) analyses.

Two reviewers (AP and VPN) independently screened by title and abstract and selected articles for full-text review and data extraction. Screening and data extraction were performed using the Covidence tool[67]. Any disagreements that arose between the reviewers were resolved through discussion.

The following variables were extracted from each study: first author, year of publication, number of participants, age, sex, musical feature, year of education, year of musical training, age of onset of musical training, training hours per week, and MRI acquisition, processing, and analysis parameters. The main outcome to extract was coordinates resulting from functional brain activity related to music perception, production and/or imagination. If any of these points were not reported in the original article, authors were contacted to retrieve this information. Nine authors were contacted, with 3 positive answers.

Quality assessment of MRI studies. A set of guidelines used for the standardization of reporting MRI results was utilized to assess the quality of the included studies[23,24]. Such guidelines dictate a more consistent and coherent policy for the reporting of MRI methods to ensure that methods can be understood and replicated.

Activation likelihood estimation (ALE) and meta-analytic connectivity modelling (MACM). All meta-analyses were performed using the activation likelihood estimation (ALE) method, implemented in GingerALE software v3.0.2[68], from the BrainMap[68] platform. Reported significant coordinates (foci) were extracted from each study. If necessary, coordinates were converted from Talairach coordinates to MNI space using the Lancaster transform (icbm2tal) incorporated in GingerALE[14,16]. The ALE method extracts the coordinates from the included studies and tests for anatomical consistency and concordance between the studies. The coordinates are weighted according to the size of the sample (number of subjects), and these weightings contribute to form estimates of anatomic likelihood estimation for each intracerebral voxel on a standardized map. This method treats anatomic foci not as single points, but as spatial probability distributions centred at the given coordinates. Thus, the algorithm tests the correlation between the spatial locations of the foci across MRI studies investigating the same construct and assesses them against a null-distribution of random spatial association between experiments[34]. Statistical significance of the ALE scores was determined by a permutation test using cluster-level inference at $p < 0.05$ (FWE), with a cluster-forming threshold set at $p < 0.001$. The primary outcome was brain functional activity related to music cognition with the aim of comprehensively examine the brain regions associated with music perception, music production, and music imagery, independently.

Then, the resulting ALE maps from each group were tested for similarity (conjunction) and difference (subtraction) in a contrast analysis with the purpose of identifying common and distinct brain areas recruited while listening, playing and/or imagining music. The GingerALE software was used following procedures described in the manual[69]. In short, during the first step, the foci from the two groups of interest were merged into a single text file and an ALE meta-analysis was conducted on the new file. Then, three ALE maps are imported to the GingerALE algorithm which randomly divides the pooled foci (foci of two datasets to be compared) into two new datasets of the same size as the original dataset. Then, voxel-wise ALE scores are calculated by subtracting one ALE image from the other. This procedure is repeated 10,000 times to yield a higher number of permutations compared to that of original dataset[70]. Finally, the 'true' difference of ALE scores is established and compared to the differences under a null distribution. This yields a voxel-wise p value of the difference which was calculated using cluster-level inference at $p < 0.05$ (FWE), with a cluster-forming threshold set at $p < 0.001$.

Meta-analytic connectivity modelling (MACM) was conducted to identify patterns of co-activation from regions-of-interest (ROI) resulting from the primary outcomes, aiming to functionally segregate each region's putative contribution to behavioural domains and paradigm classes[71,72]. Such analyses were performed using Sleuth[73] to identify studies reporting activation within each music-related ROI boundary independently and

included the experiment level search criteria of "context: normal mapping" and "activations: activation only". Then, ALE meta-analyses were performed in GingerALE over all foci resulted after the search in Sleuth to identify regions of significant convergence. Statistical significance of the ALE scores was determined by a permutation test using cluster-level inference at $p < 0.05$ (FWE), with a cluster-forming threshold set at $p < 0.001$. Functional characterization of music-related clusters was based on the "Behavioral Domain" (BD) meta-data categories in the BrianMap database which include action, perception, emotion, cognition and interoception.

All meta-analytic results (ALE maps) were visualized using Mango[21] on the MNI152 1 mm standard brain, and resulting coordinates were cross-referenced to the Harvard–Oxford Cortical and Subcortical Atlas and the Juelich Histological Atlas via NeuroVault[74] and FSLeyes[75], respectively.

Fail-Safe N analysis (FSN). Coordinate-based meta-analyses such as ALE can be subject to different forms of publication bias which may impact results and invalidate findings (e.g., the "file drawer problem"). The Fail-Safe N analysis (FSN)[64] assesses the robustness of results against potential publication bias. This method refers to the amount of contra-evidence that can be added to a meta-analysis before the results change and can be obtained for each cluster that survives thresholding in an ALE meta-analysis. It is estimated that a 95% confidence interval for the number of studies that report no local maxima varies from 5 to 30 per 100 published studies of normal human brain mapping. Using the upper bound and the fact that the CBMA's consist of 105 music perception studies, 19 music production studies, and 15 music imagery studies, an estimate for the number of unpublished experiments is 32, 6, and 5, respectively. Therefore, the minimum FSN was defined as 32 for music perception, 6 for music production, and 5 for music imagery. A higher FSN indicates more stable results and hence a higher robustness.

Data availability

The data supporting the findings of this study is freely available at the Open Science Framework (OSF) website https://osf.io/mtrha/?view_only=c85a6244897b4a7f9f96c25f9149bbcf.

Received: 20 May 2021; Accepted: 6 October 2021
Published online: 21 October 2021

References

1. Friston, K. Does predictive coding have a future?. *Nat. Neurosci.* **21**, 1019–1021 (2018).
2. Gordon, C. L., Cobb, P. R. & Id, R. B. Recruitment of the motor system during music listening: An ALE meta-analysis of fMRI data. *PLoS ONE* https://doi.org/10.1371/journal.pone.0207213 (2018).
3. Rauss, K. & Pourtois, G. What is bottom-up and what is top-down in predictive coding?. *Front. Psychol.* **4**, 276 (2013).
4. Engel, A. K., Fries, P. & Singer, W. Dynamic predictions: Oscillations and synchrony in top-down processing. *Nat. Rev. Neurosci.* **2**, 704–716 (2001).
5. Barlow, H. B. The knowledge used in vision and where it comes from. *Philos. Trans. R. Soc. B Biol. Sci.* **352**, 1141–1147 (1997).
6. Koelsch, S., Vuust, P. & Friston, K. Predictive processes and the peculiar case of music. *Trends Cogn. Sci.* **23**, 63–77 (2019).
7. Schröger, E., Kotz, S. A. & SanMiguel, I. Bridging prediction and attention in current research on perception and action. *Brain Res.* **1626**, 1–13 (2015).
8. Keller, P. E. & Keller, P. Mental imagery in music performance: Underlying mechanisms and potential benefits. *Ann. N.Y. Acad. Sci.* https://doi.org/10.1111/j.1749-6632.2011.06439.x (2012).
9. Zatorre, R. J., Chen, J. L. & Penhune, V. B. When the brain plays music: Auditory-motor interactions in music perception and production. *Nat. Rev. Neurosci.* https://doi.org/10.1038/nrn2152 (2007).
10. Sa de Almeida, J. *et al.* Music enhances structural maturation of emotional processing neural pathways in very preterm infants. *Neuroimage* **207**, 116391 (2020).
11. Kung, S.-J.S.-J., Chen, J. L., Zatorre, R. J. & Penhune, V. B. Interacting cortical and basal ganglia networks underlying findings and tapping to the musical beat. *J. Cogn. Neurosci.* **25**, 401–420 (2013).
12. Groussard, M. *et al.* When music and long-term memory interact: Effects of musical expertise on functional and structural plasticity in the hippocampus. *PLoS One* **5**, e13225 (2010).
13. Koelsch, S., Fritz, T. & Schlaug, G. Amygdala activity can be modulated by unexpected chord functions during music listening. *NeuroReport* **19**, 1815–1819 (2008).
14. Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F. & Fox, P. T. Activation likelihood estimation meta-analysis revisited. *Neuroimage* **59**, 2349–2361 (2012).
15. Carter, C. S. *et al.* Enhancing the informativeness and replicability of imaging genomics studies. *Biol. Psychiatry* **82**, 157–164 (2017).
16. Laird, A. R. *et al.* Comparison of the disparity between Talairach and MNI coordinates in functional neuroimaging data: Validation of the Lancaster transform. *Neuroimage* https://doi.org/10.1016/j.neuroimage.2010.02.048 (2010).
17. Wager, T. D., Lindquist, M. A., Nichols, T. E., Kober, H. & Van Snellenberg, J. X. Evaluating the consistency and specificity of neuroimaging data using meta-analysis. *Neuroimage* **45**, S210–S221 (2009).
18. Chauvigné, L. A. S., Gitau, K. M. & Brown, S. The neural basis of audiomotor entrainment: An ALE meta-analysis. *Front. Hum. Neurosci.* **8**, 776 (2014).
19. Samson, F., Zeffiro, T. A., Toussaint, A. & Belin, P. Stimulus complexity and categorical effects in human auditory cortex: An activation likelihood estimation meta-analysis. *Front. Psychol.* **1**, 241 (2011).
20. Koelsch, S. A coordinate-based meta-analysis of music-evoked emotions. *Neuroimage* **223**, 117350 (2020).
21. Lancaster, J. L. *et al.* Automated regional behavioral analysis for human brain images. *Front. Neuroinform.* **6**, 1–12 (2012).
22. Choudhury, S., Fishman, J. R., McGowan, M. L. & Juengst, E. T. Big data, open science and the brain: Lessons learned from genomics. *Front. Hum. Neurosci.* **8**, 239 (2014).
23. Poldrack, R. A. *et al.* Guidelines for reporting an fMRI study. *Neuroimage* **40**, 409–414 (2008).
24. Nichols, T. E. *et al.* Best practices in data analysis and sharing in neuroimaging using MRI. *Nat. Neurosci.* **20**, 299–303 (2017).
25. Saygin, A. P., Leech, R. & Dick, F. Nonverbal auditory agnosia with lesion to Wernicke's area. *Neuropsychologia* **48**, 107–113 (2010).
26. Koelsch, S. Neural substrates of processing syntax and semantics in music. in *Music that Works: Contributions of Biology, Neurophysiology, Psychology, Sociology, Medicine and Musicology* 143–153. https://doi.org/10.1007/978-3-211-75121-3_9 (2009).
27. Grahn, J. A. Neural mechanisms of rhythm perception: Current findings and future perspectives. *Top. Cogn. Sci.* https://doi.org/10.1111/j.1756-8765.2012.01213.x (2012).

28. Edagawa, K. & Kawasaki, M. Beta phase synchronization in the frontal-temporal-cerebellar network during auditory-to-motor rhythm learning. *Sci. Rep.* **7**, 1–9 (2017).
29. Chen, J. L., Penhune, V. B. & Zatorre, R. J. Listening to musical rhythms recruits motor regions of the brain. *Cereb. Cortex* **18**, 2844–2854 (2008).
30. Snyder, J. S. & Large, E. W. Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cogn. Brain Res.* **24**, 117–126 (2005).
31. Zanto, T. P., Snyder, J. S. & Large, E. W. Neural correlates of rhythmic expectancy. *Adv. Cogn. Psychol.* **2**, 221–231 (2006).
32. Meck, W. H., Penney, T. B. & Pouthas, V. Cortico-striatal representation of time in animals and humans. *Curr. Opin. Neurobiol.* **18**, 145–152 (2008).
33. Karmarkar, U. R. & Buonomano, D. V. Timing in the absence of clocks: Encoding time in neural network states. *Neuron* **53**, 427–438 (2007).
34. Chen, J. L., Penhune, V. B. & Zatorre, R. J. Moving on time: Brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *J. Cogn. Neurosci.* **20**, 226–239 (2008).
35. Graybiel, A. M. Building action repertoires: Memory and learning functions of the basal ganglia. *Curr. Opin. Neurobiol.* **5**, 733–741 (1995).
36. Butti, C. & Hof, P. R. The insular cortex: A comparative perspective. *Brain Struct. Funct.* **214**, 477–493 (2010).
37. Guenot, M., Isnard, J. & Sindou, M. Surgical anatomy of the insula. *Adv. Tech. Stand. Neurosurg.* **29**, 265–288 (2004).
38. Shura, R. D., Hurley, R. A. & Taber, K. H. Insular cortex: Structural and functional neuroanatomy. *J. Neuropsychiatry Clin. Neurosci.* **26**, 277–282. https://doi.org/10.1176/appi.neuropsych.260401 (2014).
39. Gogolla, N. The insular cortex. *Curr. Biol.* **27**, R580–R586 (2017).
40. Kurth, F., Zilles, K., Fox, P. T., Laird, A. R. & Eickhoff, S. B. A link between the systems: Functional differentiation and integration within the human insula revealed by meta-analysis. *Brain Struct. Funct.* https://doi.org/10.1007/s00429-010-0255-z (2010).
41. Bamiou, D. E., Musiek, F. E. & Luxon, L. M. The insula (Island of Reil) and its role in auditory processing: Literature review. *Brain Res. Rev.* **42**, 143–154 (2003).
42. Craig, A. D. Emotional moments across time: A possible neural basis for time perception in the anterior insula. *Philos. Trans. R. Soc. B* https://doi.org/10.1098/rstb.2009.0008 (2009).
43. Koelsch, S. Towards a neural basis of music-evoked emotions. *Trends Cogn. Sci.* **14**, 131–137 (2010).
44. Hoshi, E. & Tanji, J. Functional specialization in dorsal and ventral premotor areas. *Prog. Brain Res.* **143**, 507–511 (2004).
45. Króliczak, G., Piper, B. J. & Frey, S. H. Specialization of the left supramarginal gyrus for hand-independent praxis representation is not related to hand dominance. *Neuropsychologia* **93**, 501–512 (2016).
46. Chen, J. L., Zatorre, R. J. & Penhune, V. B. Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms. *Neuroimage* **32**, 1771–1781 (2006).
47. Morecraft, R. J. *et al.* Cytoarchitecture and cortical connections of the anterior cingulate and adjacent somatomotor fields in the rhesus monkey. *Brain Res. Bull.* **87**, 457–497 (2012).
48. Herrojo Ruiz, M., Maess, B., Altenmüller, E., Curio, G. & Nikulin, V. V. Cingulate and cerebellar beta oscillations are engaged in the acquisition of auditory-motor sequences. *Hum. Brain Mapp.* **38**, 5161–5179 (2017).
49. Satoh, M., Takeda, K., Nagata, K., Hatazawa, J. & Kuzuhara, S. Activated brain regions in musicians during an ensemble: A PET study. *Cogn. Brain Res.* **12**, 101–108 (2001).
50. Brodal, P., Bjaalie, J. G. & Aas, J. E. Organization of cingulo-ponto-cerebellar connections in the cat. *Anat. Embryol. (Berl)* **184**, 245–254 (1991).
51. Royal, I. *et al.* Activation in the right inferior parietal lobule reflects the representation of musical structure beyond simple pitch discrimination. *PLoS One* **11**, e0155291 (2016).
52. Callow, N., Roberts, R., Hardy, L., Jiang, D. & Edwards, M. G. Performance improvements from imagery: Evidence that internal visual imagery is superior to external visual imagery for slalom performance. *Front. Hum. Neurosci.* **7**, 697 (2013).
53. Raij, T. T. & Riekki, T. J. J. Poor supplementary motor area activation differentiates auditory verbal hallucination from imagining the hallucination. *NeuroImage Clin.* **1**, 75–80 (2012).
54. Bareham, C. A. *et al.* Role of the right inferior parietal cortex in auditory selective attention: An rTMS study. *Cortex* **99**, 30–38 (2018).
55. Koenigs, M., Barbey, A. K., Postle, B. R. & Grafman, J. Superior parietal cortex is critical for the manipulation of information in working memory. *J. Neurosci.* **29**, 14980–14986 (2009).
56. Kraeutner, S. N., Keeler, L. T. & Boe, S. G. Motor imagery-based skill acquisition disrupted following rTMS of the inferior parietal lobule. *Exp. Brain Res.* **234**, 397–407 (2016).
57. Guillot, A. *et al.* Brain activity during visual versus kinesthetic imagery: An fMRI study. *Hum. Brain Mapp.* **30**, 2157–2172 (2009).
58. Seghier, M. L. The angular gyrus: Multiple functions and multiple subdivisions. *Neuroscientist* **19**, 43–61 (2013).
59. Emch, M., von Bastian, C. C. & Koch, K. Neural correlates of verbal working memory: An fMRI meta-analysis. *Front. Hum. Neurosci.* **13**, 180 (2019).
60. de Boer, D. M. L., Johnston, P. J., Kerr, G., Meinzer, M. & Cleeremans, A. A causal role for the right angular gyrus in self-location mediated perspective taking. *Sci. Rep.* **10**, 1–10 (2020).
61. Schubotz, R. I., Friederici, A. D. & Yves Von Cramon, D. Time perception and motor timing: A common cortical and subcortical basis revealed by fMRI. *Neuroimage* **11**, 1–12 (2000).
62. Shomstein, S. Cognitive functions of the posterior parietal cortex: Top-down and bottom-up attentional control. *Front. Integr. Neurosci.* **6**, 38 (2012).
63. Maes, P.-J., Leman, M., Palmer, C. & Wanderley, M. Action-based effects on music perception. *Front. Psychol.* **4**, 1008 (2014).
64. Acar, F., Seurinck, R., Eickhoff, S. B. & Moerkerke, B. Assessing robustness against potential publication bias in Activation Likelihood Estimation (ALE) meta-analyses for fMRI. *PLoS One* **13**, e0208177 (2018).
65. Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., Welch, V. A. (editors). *Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021)*. Available from http://www.training.cochrane.org/handbook (Cochrane, 2021).
66. Moher, D., Liberati, A., Tetzlaff, J. & Altman, D. G. Preferred reporting items for systematic reviews and meta-analyses. *The PRISMA Statement* **6**, 1–5 (2009).
67. Covidence systematic review software, Veritas Health Innovation, Melbourne, Australia. http://www.covidence.org. https://support.covidence.org/help/how-can-i-cite-covidence
68. Eickhoff, S. B. *et al.* Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. *Hum. Brain Mapp.* https://doi.org/10.1002/hbm.20718 (2009).
69. Fox, P. T. *et al. User Manual for GingerALE 2.3*.
70. Alain, C., Du, Y., Bernstein, L. J., Barten, T. & Banai, K. Listening under difficult conditions: An activation likelihood estimation meta-analysis. *Hum. Brain Mapp.* **39**, 2695–2709 (2018).
71. Laird, A. R. *et al.* Networks of task co-activations. *Neuroimage* **80**, 505–514 (2013).
72. Robinson, J. L., Laird, A. R., Glahn, D. C., Lovallo, W. R. & Fox, P. T. Metaanalytic connectivity modeling: Delineating the functional connectivity of the human amygdala. *Hum. Brain Mapp.* **31**, 173–184 (2010).

73. Laird, A. R., Lancaster, J. L. & Fox, P. T. BrainMap: The social evolution of a human brain mapping database. *Neuroinformatics* **3**, 65–78 (2005).
74. Gorgolewski, K. J. *et al.* NeuroVault.org: A web-based repository for collecting and sharing unthresholded statistical maps of the human brain. *Front. Neuroinform.* **9**, 1–9 (2015).
75. Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W. & Smith, S. M. FSL. *Neuroimage* **62**, 782–790 (2012).

Acknowledgements

The Center for Music in the Brain (MIB) is supported by The Danish National Research Foundation (grant number DNRF 117). L.B. is supported by the Carlsberg Foundation (project number CF20-0239) and by the Center for Music in the Brain. The authors wish to thank Hella Kastbjerg for assistance with language check and proofreading.

Author contributions

V.P.N. and A.P. designed the meta-analysis. V.P.N. and A.P. conducted the screening of the studies and data extraction. L.B. performed the sensitivity analysis. V.P.N. and A.P. wrote the first draft of the manuscript and prepared the initial versions of the figures and tables. L.B. edited and wrote paragraphs in the "Introduction" and "Discussion". V.P.N. prepared the final figures and tables. P.V. contributed to financially support the study and commented the final versions of the manuscript. V.P.N. finalized the paper for submission and all authors edited the manuscript and approved its final version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-00139-3.

Correspondence and requests for materials should be addressed to V.P.-N.

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

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