Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Behavioral Sciences

Music and big data: a new frontier

David M Greenberg[1,2] and Peter J Rentfrow[3]

There is an unprecedented opportunity for psychologists and behavioral scientists to merge prior theory and research with big data to develop profound insights into the way people use and are affected by music. There are now **streaming services** that store data from millions of people on their day-to-day musical listening habits; **song-level** data that tags sonic and emotion attributes for millions of songs; **wearable devices** (e.g. watches and earbuds) that capture physiological metrics including heartrate and galvanic skin response; **mobile technologies** that track a person's moment-to-moment activity, location, mood, and sociability; and **survey instruments** and **digital footprints** that capture personality and other biopsychosocial metrics in just under a minute. We propose that merging these technologies can create a new age in music psychology that exponentially expands the present knowledge and scope of the field. The new data will advance general areas of music psychology, but will also provide an important opportunity to establish new knowledge about health and well-being that can have a direct impact on the public. By scientifically mapping how music changes behavior and health in the short-term and long-term, Big Music Data can lead to future health initiatives including the development of new evidence-based treatment modalities to be utilized by medical physicians and mental health practitioners. Importantly, industry and streaming services can use these new insights to optimize their technologies and develop music-based health and wellness platforms aimed at improving the well-being of its users, ultimately impacting the way music is used by millions of people globally.

Addresses
[1] Department of Clinical Psychology, City College of New York, City University of New York, USA
[2] Department of Psychiatry, University of Cambridge, UK
[3] Department of Psychology, University of Cambridge, UK

Corresponding author: Greenberg, David M (dmg39@cam.ac.uk)

Current Opinion in Behavioral Sciences 2017, **18**:50–56

This review comes from a themed issue on **Big data in the behavioural sciences**

Edited by **Michal Kosinski** and **Tara Behrend**

http://dx.doi.org/10.1016/j.cobeha.2017.07.007

Introduction

Music is pre-historic and its impact on health has been documented for thousands of years [1,2•,3]. Music is not simply entertainment — scientific research from the past decade has shown that it played an integral role in human evolution, and is closely tied to communication, social bonding, and human development [4–8]. Today, music remains a central part of human experience across cultures and the lifespan [9]. People listen to it in a multitude of contexts for a total of 11–44% of their waking lives [9–13]. With the advancement of big data and technology, we are entering a unique period with an opportunity to gain new insights into the uses and effects of music that was not previously possible. This new knowledge has the potential to scientifically inform how music can benefit medical and clinical settings, and industry. In this paper, we will review prior scientific research on music, explore different areas of Big Music Data, and outline a conceptual approach that we feel will be most fruitful this endeavor.

Prior research

Prior theory and research in music psychology has shown the impact of music on neurological, personological, social, and cultural levels: Research in music cognition has outlined how we process musical information and form cognitive and affective representations of it [14]; research in neuroscience has identified the reward networks in the brain that are activated when listening to music and the hormones and chemicals that are excreted in response to it [15–17]; research on affect and music has explored the underlying mechanisms for how people perceive and experience emotions from music and the process of how the emotional intensions of the musician are expressed through music and then perceived and felt by the listener [18,19]; research in personality and social psychology has shown how individual differences in musical preferences is linked to personality, values, and cognitive styles [20,21,22•,23]; research on the uses and gratifications of music has shown that music plays distinct functions in a variety of contexts including concentration during work and motivation during fitness [24,12]; medical research has shown how music can impact physical health and rehabilitation such as increasing recovery rates after surgery [25•]; and research in music therapy has shown that music-based treatment interventions can be successfully used to address mental and emotional health issues in those with autism, depression, post-traumatic stress disorder, and dementia [26,27•,28,29•]. Findings in the areas have provided a thorough theoretical and scientific base which can be immediately applied to big data research.

Limitations of prior research

Though there have been significant advances in our scientific understanding of music, the field has been hindered by methodological limitations. First, with the exception of a few studies [9,30], the majority of research in music psychology has reported relatively small samples sizes. This is in part due to limited access to undergraduate samples and subject pools, less funding opportunities, and lower research exposure than is available for research in more traditional fields. Though in recent years online recruitment platforms have increased the scope of recruitment strategies, the sample sizes reported in music psychology journals remain smaller when compared to related fields like personality and social psychology. Larger samples are important because they increase statistical power, and allow researchers to control for confounding variables, observe small effects, analyze within and across sample replication, and examine affects across age groups and geography. Considering that small sample sizes contributed to lack of replication in social psychology, larger samples could prevent such a crisis in music psychology.

Second, the representative musical stimuli used in studies are limited. Often, experimental and correlational studies use stimuli that are brief (15–30 s in length) sometimes it is computer generated or manipulated, and in some cases has never been heard before by participants [31,32]. Though these approaches limit confounding variables, they also lack ecological validity. Third, regardless of the stimuli administered, given the time restrains of online and laboratory experiments, the music that is used does not capture the ways in which people naturally interact with music in their daily life, and the breadth of music they are exposed and listen to — it only captures a snapshot of the way someone listens to, responds to, and engages with music. As will be shown in this paper, big data has the possibility to advance beyond these limitations.

Big Music Data

Recent technological advances, including the Internet, streaming services, online social media, and audio file formats, has generated the collection of large amounts of data relevant for psychological research on music. This includes big data on the human-level and song-level. To date, there have been four approaches to Big Music Data discussed below.

Mass Internet surveys

One approach to Big Music Data is administering music surveys and experiments to masses of individuals. This predominantly descriptive and correlational approach provides a powerful platform for mapping a range of psychological and music-related phenomena. Typically, the data generated by online surveys are self-reports of demographic and psychological characteristics combined with information about music use that includes musical preferences, and affective reactions or perceptions of musical stimuli. The primary advantage of this approach is that it provides access to large and diverse samples of people around the world.

There have been several examples of successful platforms where large amounts of music psychology data have been collected: (a) In the myPersonality project [33•], over 20,000 Facebook users provided their affective responses to musical stimuli and completed measures of personality and other psycho-demographic measures. From this data, researchers have been able to examine the structure of musical preference and its correlates with personality and cognitive styles [21,22•,31,34]; (b) As part of the Internet-based music preferences project at Out of Service (www.outofservice.com/music-personality-test), a quarter of a million participants have provided self-reports of their personalities, demographics, and music genre-preferences. From this dataset, researchers have been able to examine musical preferences across the lifespan, showing that normative trends in musical preferences correspond to Erikson's psychosocial stages of development [9]; (c) As part of the BBC 'Lab UK' project, nearly 150,000 participants completed self-report and behavioral tasks of musical ability. These data were used by researchers to explore the structure, correlates, and geographic distribution of musical sophistication [30]. Further, the data were combined with a separate BBC 'Lab UK' dataset on personality and well-being to show that personality traits predict musical ability in both musicians and non-musicians [35]; (d) Most recently, Greenberg created the Musical Universe project (www.musicaluniverse.org), which is one of the most extensive datasets in terms of the quantity and breadth of musical and psychological variables. Over 100,000 people have completed measures on musical preferences, personality, well-being, and demographics (including musicianship, musical consumption, music training, geographic location and clinical diagnoses). In terms of psychological information, large subsamples have completed measures on mood, cognitive style, emotion regulation, values, the dark triad of personality, and mind-reading. In terms musical information, large subsamples have completed measures on musical engagement style, music perception, musical and creative arts performance attributes, and experimental listening tasks that detected changes empathy in response to music listening.

Online social media

Online social media (OSM) is one area where musical behavior can be observed 'in the wild' via digital footprints. OSM are forums where people come together for the purpose of interacting with each other and sharing information — an interaction which invariably includes the expression of musical information. Typically, the data available from OSM include behavioral records of the music people like that can be gained from Facebook likes and Twitter. One of the biggest advantages of these data is that they are behavioral and therefore overcome some of the limitations of self-report methods. This information can supplement mass Internet surveys insofar as participants can be recruited from social media to complete a survey which they can then share with friends.

There are several ways researchers can freely access OSM data. For example, a multitude of research teams across disciplines have turned to Twitter (a micro-blogging site where people can send and receive short messages) to address research questions. Data from Twitter can be gathered in various ways. First, their flexible terms of use policy allow researchers to use publically posted tweets and information in their research. Second, Twitter previously developed an API where researchers can download a small percentage (e.g. 1%) of their data. And third, Twitter has made portions of their data (upwards to 10%) available for researchers in the past. Psychologists have made use of this data in a variety of ways, including understanding how language on OSM is linked to personality and wellbeing on both individual and geographic levels [36,37]. This can be done by using linguistic analysis (e.g. NLP) of tweets combine with self-reports of personality [38].

Twitter data also provides an opportunity for those researching music. Computer scientists created the freely available #nowplaying music dataset (http://dbis-nowplaying.uibk.ac.at) [39]. This dataset includes music listening behavior and events of posted by Twitter users including metadata consisting of artist and track information. The dataset is continually updated and currently includes 62,217,458 tweets; 2,296,758 users; 1,507,084 tracks; and 172,313 artists.

There are other avenues in which to use digital footprints including Facebook activity and likes. Facebook likes can be isolated to artists, tracks, and styles as a way of capturing music preference data. These music-specific Facebook likes can be linked to a host of psychological variables and mapped across friendship dyads and networks. A pre-existing large dataset of Facebook likes is freely available to researchers at www.mypersonality.org.

The disadvantage of OSM data is that it does not actually capture musical listening behaviors but rather the thoughts, beliefs, and attitudes about them. Further, these thoughts, beliefs, and attitudes displayed on social media are likely to be biased due to social desirability and the social connotations tied to musical behavior. Prior research has shown that people use music as an identity badge [40], and that people make judgements of others based on their musical preferences — that these judgements are both agreed upon and contain a kernel of truth [41,42]. Therefore, people's music-specific digital footprints on social media platforms are more likely to be influenced my social connotations and meeting social goals than other big data platforms.

Music streaming services

The platforms with the most upside and ecological validity is data from music streaming services that include logs of people's daily music listening behaviors. Popular music streaming services include Spotify, Apple Music, YouTube, Pandora, and Last.FM. Not only can this data capture musical preferences with ecological validity through the listening histories of its users over the course of years, but it can also capture users' listening habits (e.g. whether users listen to songs on repeat, in full, or repeats or skips certain sections of songs) [43,44]. Such data have the most promise because they truly capture listening behavior 'in the wild', consist of a large proportion of the users' listening behavior that they engage with in their daily lives, record listening histories over long periods of time, and have the potential to be synchronized with other big data methods and technologies. However, the disadvantage of music streaming data alone is that it does not capture in-depth psychological characteristics of users, which makes it difficult to study the effects that music has on listeners. Using streaming data may be most challenging since many streaming services do not allow researchers access to their data (see Last.FM as an exception here: www.mypersonality.org). This is one of the areas where collaboration between industry and academia can be most fruitful.

Song databases

In addition to Big Music Data on the person-level, there is also Big Music Data on the song-level. This includes meta-data on the genre labels, tags, and perceived attributes of songs. This data was originally developed from human raters coding music for different attributes like in the Music Genome Project, however, with advances in machine learning (ML), there are now several ML-based software programs that can detect and extract features in songs. For example, a recent study examined the variation of musical attributes in 17,000 billboard songs from the past half-century [45]. Many of these databases have been generated and utilized for industry and commercial purposes and music applications such as Shazaam that can recognize a song based on listening to it for just a few seconds. Examples of these databases include the Million Song Dataset, and companies like Gracenote and The Echo Nest (now part of Spotfiy).

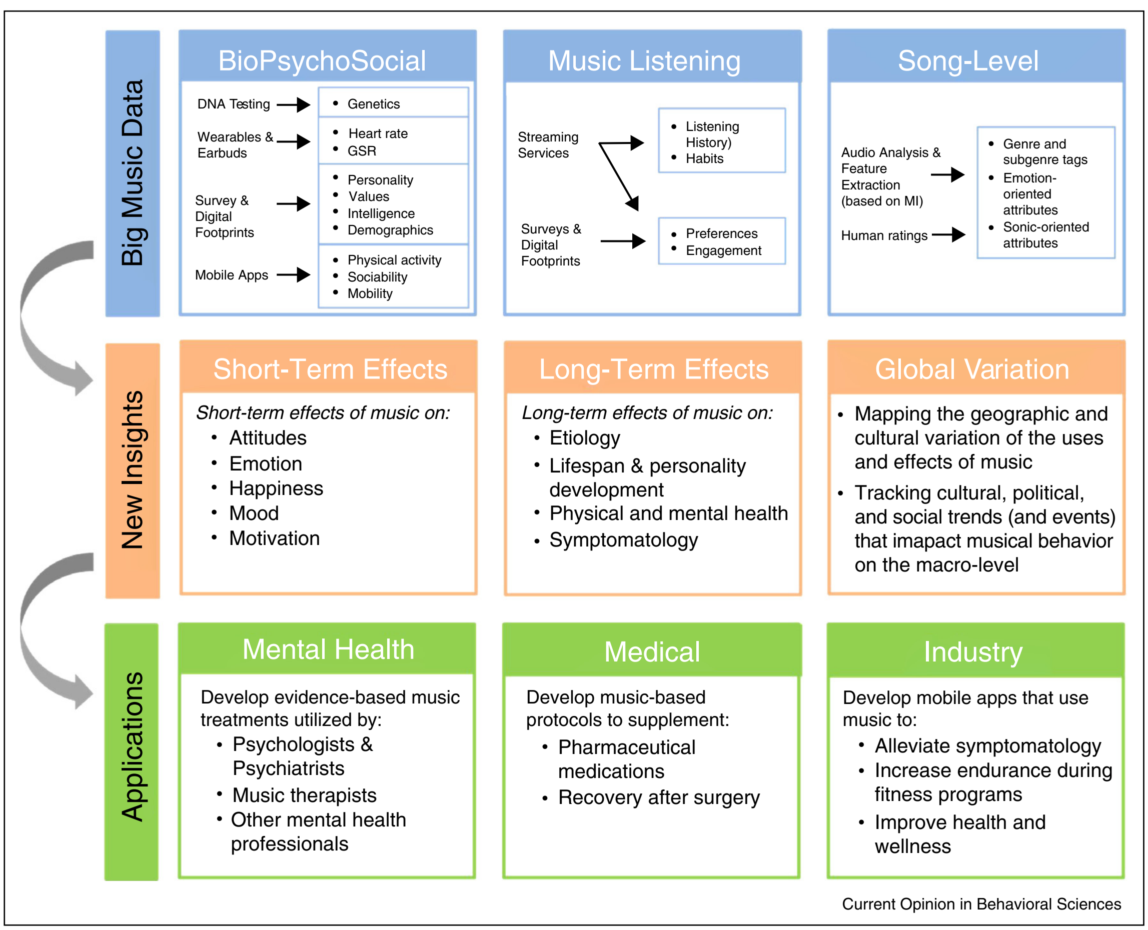
There are alternative routes to examining song-level data. SoundCloud may be an interesting resource considering that music on SoundCloud has tags in addition to comments (e.g. emotions or climax in a song) that are marked at different time points on each track by the listeners. The latter provides insight into human perception of music and if aggregated across tracks can provide maps into how song-level sonic and emotional attributes link to human judgment and perception of the music.

To date, song-level data have been used primarily by computer scientists and only recently has begun to be used by psychologists. For example, a recent study [22•] found that perceived music attributes at the song-level are organized in a way that reflects previously established psychological models of affect including the circumplex model of affect [46] and the positive and negative affect framework used to conceptualize mood states [47]. However, a more detailed understanding of song-level music attributes is needed. Understanding how the specific harmonic, rhythmic, melodic, and dynamic structures elicit psychological and emotional responses in the listener. ML and AI techniques will be particularly useful in this endeavor to understand the connections between musical structure and human response on a large scale.

Big opportunities

Both theory and research suggest that people use and are affected by music in a way that fulfills their psychological needs [48]. However, in large part because of laboratory and time constraints, understanding the uses and effects of music has been limited to artificial settings. And though neuroscience has shown how music listening is linked to brain structure and activity, small samples of participants that have differing preferences and prior musical experience have limited generalizability. In contrast, Big Music Data presents the opportunity to map the intricacies of musical behavior on biological, personological, social, and cultural levels. Most importantly, it presents an unprecedented opportunity to understand how music can change behavior and impact health and wellness. To gain this new understanding, psychologists and behavioral scientists will have to take an approach that combines information about biopsychosocial variables with both music listening histories and song-level meta-data. With Big Music Data, there is the possibility observe a systematic constellation of biopsychosocial and music variables that lead to behavior change and improvements in health and wellness. The knowledge can then be used to develop medical, clinical, and industry applications aimed at the psychological needs of an individual and improving health and wellness. A conceptual map is visually displayed in Figure 1 and elaborated below.

Figure 1



Conceptual Map of Big Music Data, new insights to be gained, and applications to be developed.

Merging technologies

There is a plethora of data from new technologies that is being collected daily. Physiological variables like heart rate and galvanic skin responses (GSR) can now be gained from wearable devices including watches and earbuds. Psycho-demographic profiles that include personality, values, intelligence, and demographics (sex, age, socio-economic status, ethnicity, clinical diagnoses) can be measured with validated self-reports instruments that take less than a minute to complete (e.g. TIPI: [49]), or through digital footprints which can be accessed immediately from social media activity (e.g. www.applymagicsauce.com: [33•]). Moment-to-moment physical activity, sociability, and mobility can be measured from mobile apps like EmotionSense (www.emotionsense.org: [50•]). These are all metrics that are pre-existing or require little effort to complete. Further upon the horizon are possibilities to include Big Data from genetics. Several genomic biotechnology companies (e.g. 23andMe) have collected additional data from their customers via surveys that include scientific measures (e.g. on personality and clinical diagnoses) and have teamed with scientists to publish findings [51•]. These approaches and technologies (particularly those that continually collect data such as physiological metrics) combined with exhaustive listening histories from streaming services (that include periodic longitudinal assessments that track behavior, clinical diagnoses) and song-level meta-data, will provide a robust framework for insights to be developed about music. Merging these data and technologies together will require much greater collaboration and communication between experts across disciplines (e.g. psychologists, data and computer scientists, geneticists, brain imagers, musicians, music therapists, and industry) than the current status quo in music psychology.

New insights

By using technologies that continually track behavior (e.g. digital footprints and mobile apps) in conjunction with surveys, insights into the short and long-term effects of music will be feasible. Short-term effects include behavior change such as alterations in attitude, emotion, happiness, mood, and motivation. With streaming data that longitudinally tracks listening history and habits (not just months but potentially decades) researchers can observe how music is intertwined with the etiology of clinical diagnoses, the alleviation of symptomology, general physical and mental health, and lifespan and personality development.

The combined approach will also allow new insights to be gained about the uses and effects of music on a global scale. One of the most assumed notions about music is that it is a universal language. Though some studies suggest that music is perceived similarly across cultures, others suggest that there are important differences [52]. With Big Music Data, it will be possible to map music preferences, engagement, and performance attributes globally. Linking this information to individual level data on personality and well-being in addition to culture variables like population density, climate, national character and geographic location, will enable valuable insights about how music and culture are intertwined. Moreover, from streaming data it will be possible to observe how national, cultural, political, and social trends and events impact musical behavior globally.

Applications

These new insights can be used as the foundation of data-driven applications for health professionals and industry. Knowledge about how music listening links to the diagnostic etiology and the alleviation of symptoms, can be used to develop evidence-based treatments targeted at different populations (e.g. depression, autism, PTSD). This knowledge can be used by a variety of mental health professionals including clinical psychologists, music therapists, and psychiatrists. Findings from Big Music Data can also be used to develop music-based protocols that can supplement pharmaceutical medications, or to be used in hospitals pre-operation and post-operation to boost recovery [25•]. Importantly, the new insights can be used by industry and streaming services to develop health and wellness apps for their users aimed at increasing fitness and well-being.

Conclusion

There are fantastic opportunities forthcoming with Big Data that can be a new frontier in generating knowledge and technology on the beneficial psychological powers of music. Here we have highlighted the ways in which this new knowledge and technology can be used by industry and clinicians to have a positive impact on individuals and society. However, we caution that such information in the wrong hands can also be used to promote negative behaviors and increase aggression, distractibility, and addictive behavior. Manipulation of lyrics and other musical elements can increase negative outgroup sentiments just as much as positive sentiments. It can cause division just as much as cohesion. These are ethical considerations that must be at the forefront in the minds of psychologists and behavioral scientists when conducting this research.

Theory and research from the past decades has overwhelming shown how music uses and affects are inextricably biological, psychological, and social in nature. Yet, some still view music research as unimportant. As reviewed in this paper, the impact music has on human life in profound. With the emergence of Big Music Data, there will be opportunities for researchers to generate 'knowledge of knowledge's sake' and to publish flashy findings that will undoubtedly attract media attention. But we urge researchers to concentrate their efforts on conducting research with Big Music Data that will have real world implications that will ultimately benefit individuals and society.

Conflict of interest statement

Nothing declared.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest

1. Blacking J: *Music, Culture, and Experience: Selected Papers of John Blacking*. London: University of Chicago Press; 1995.
2. • Conrad C: **Music for healing: from magic to medicine**. *Lancet* 2010, **376**:1980-1981.

A terrific and succinct overview of various historical roots in music and healing and how it can inform contemporary work on music research and therapy.

3. Higham T, Basell L, Jacobi R, Wood R, Ramsey CB, Conard NJ: **Testing models for the beginnings of the Aurignacian and the advent of figurative art and music: the radiocarbon chronology of Geißenklösterle**. *J Hum Evol* 2012, **62**:664-676.
4. Cross I, Morley I: **The evolution of music: theories, definitions and the nature of the evidence**. In *Communicative Musicality.* Edited by Malloch S, Trevarthen C. Oxford: Oxford University Press; 2008:61-82.
5. Brown S: **The 'musilanguage' model of music evolution**. In *The Origins of Music.* Edited by Wallin N, Merker B, Brown S. Cambridge, MA: MIT Press; 2000:271-300.
6. Huron D: **Is music an evolutionary adaptation?** *Ann N Y Acad Sci* 2001, **930**:43-61.
7. Loersch C, Arbuckle NL: **Unraveling the mystery of music: music as an evolved group process**. *J Personal Soc Psychol* 2013, **105**:777.
8. Mithen S: *The Singing Neanderthals: The Origins of Music, Language, Mind and Body*. London: Weidenfeld & Nicolson; 2005.
9. Bonneville-Roussy A, Rentfrow PJ, Xu MK, Potter J: **Music through the ages: trends in musical engagement and preferences from adolescence through middle adulthood**. *J Personal Soc Psychol* 2013, **105**:703.
10. Juslin PN, Liljeström S, Västfjäll D, Barradas G, Silva A: **An experience sampling study of emotional reactions to music: listener, music, and situation**. *Emotion* 2008, **8**:668.
11. Motion Picture Association of America, Inc.: *Entertainment Industry Market Statistics [Online]*. 2007. Retrieved on 15 September 2011 from: http://www.mpaa.org/USEntertainmentIndustryMarketStats.pdf.
12. North AC, Hargreaves DJ, Hargreaves JJ: **Uses of music in everyday life**. *Music Percept: Interdiscip J* 2004, **22**:41-77.
13. Sloboda JA, O'Neill SA, Ivaldi A: **Functions of music in everyday life: an exploratory study using the Experience Sampling Method**. *Music Sci* 2001, **5**:9-32.
14. Patel AD: **Language, music, syntax and the brain**. *Nat Neurosci* 2003, **6**:674-681.
15. Blood AJ, Zatorre RJ: **Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion**. *Proc Natl Acad Sci U S A* 2001, **98**:11818-11823.
16. Huron D: **Why is sad music pleasurable? A possible role for prolactin**. *Music Sci* 2011, **15**:146-158.
17. Menon V, Levitin DJ: **The rewards of music listening: response and physiological connectivity of the mesolimbic system**. *Neuroimage* 2005, **28**:175-184.
18. Juslin PN: **Cue utilization in communication of emotion in music performance: relating performance to perception**. *J Exp Psychol: Hum Percept Perform* 2000, **26**:1797.
19. Juslin PN, Västfjäll D: **Emotional responses to music: the need to consider underlying mechanisms**. *Behav Brain Sci* 2008, **31**:559-575.
20. Boer D, Fischer R, Strack M, Bond MH, Lo E, Lam J: **How shared preferences in music create bonds between people: values as the missing link**. *Personal Soc Psychol Bull* 2011. 0146167211407521.
21. Greenberg DM, Baron-Cohen S, Stillwell DJ, Kosinski M, Rentfrow PJ: **Musical preferences are linked to cognitive styles**. *PLOS ONE* 2015, **10**:e0131151.
22. • Greenberg DM, Kosinski M, Stillwell DJ, Monteiro BL, Levitin DJ, Rentfrow PJ: **The song is you: preferences for musical attribute dimensions reflect personality**. *Soc Psychol Personal Sci* 2016, **7**:597-605.
23. Rentfrow PJ, Gosling SD: **The do re mi's of everyday life: the structure and personality correlates of music preferences**. *J Personal Soc Psychol* 2003, **84**:1236.
24. Lonsdale AJ, North AC: **Why do we listen to music? A uses and gratifications analysis**. *Br J Psychol* 2011, **102**:108-134.
25. • Hole J, Hirsch M, Ball E, Meads C: **Music as an aid for postoperative recovery in adults: a systematic review and meta-analysis**. *Lancet* 2015, **386**:1659-1671.
26. Carr C, d'Ardenne P, Sloboda A, Scott C, Wang D, Priebe S: **Group music therapy for patients with persistent post-traumatic stress disorder – an exploratory randomized controlled trial with mixed methods evaluation**. *Psychol Psychother: Theory Res Pract* 2012, **85**:179-202.
27. • Gold C, Solli HP, Krüger V, Lie SA: **Dose–response relationship in music therapy for people with serious mental disorders: systematic review and meta-analysis**. *Clin Psychol Rev* 2009, **29**:193-207.
28. Ueda T, Suzukamo Y, Sato M, Izumi SI: **Effects of music therapy on behavioral and psychological symptoms of dementia: a systematic review and meta-analysis**. *Age Res Rev* 2013, **12**:628-641.
29. • Whipple J: **Music in intervention for children and adolescents with autism: a meta-analysis**. *J Music Therapy* 2004, **41**:90-106.
30. Müllensiefen D, Gingras B, Musil J, Stewart L: **The musicality of non-musicians: an index for assessing musical sophistication in the general population**. *PLOS ONE* 2014, **9**:e89642.
31. Rentfrow PJ, Goldberg LR, Levitin DJ: **The structure of musical preferences: a five-factor model**. *J Personal Soc Psychol* 2011, **100**:1139.
32. Vuoskoski JK, Thompson WF, McIlwain D, Eerola T: **Who enjoys listening to sad music and why?** *Music Percept: Interdiscip J* 2012, **29**:311-317.
33. • Kosinski M, Stillwell D, Graepel T: **Private traits and attributes are predictable from digital records of human behavior**. *Proc Natl Acad Sci* 2013, **110**:5802-5805.
34. Rentfrow PJ, Goldberg LR, Stillwell DJ, Kosinski M, Gosling SD, Levitin DJ: **The song remains the same: a replication and extension of the MUSIC model**. *Music Percept: Interdiscip J* 2012, **30**:161-185.
35. Greenberg DM, Müllensiefen D, Lamb ME, Rentfrow PJ: **Personality predicts musical sophistication**. *J Res Personal* 2015, **58**:154-158.
36. Eichstaedt JC, Schwartz HA, Kern ML, Park G, Labarthe DR, Merchant RM, Jha S, Agrawal M, Dziurzynski LA, Sap M, Weeg C: **Psychological language on Twitter predicts county-level heart disease mortality**. *Psychol Sci* 2015, **26**:159-169.
37. Quercia D, Kosinski M, Stillwell D, Crowcroft J: **Our Twitter profiles, our selves: predicting personality with Twitter**. *2011 IEEE International Conference on Privacy, Security, Risk, and Trust, and IEEE International Conference on Social Computing*; *IEEE, Piscataway, NJ: 2011:180-185*.
38. Schwartz HA, Eichstaedt JC, Kern ML, Dziurzynski L, Ramones SM, Agrawal M, Shah A, Kosinski M, Stillwell D, Seligman ME, Ungar LH: **Personality, gender, and age in the**

language of social media: the open-vocabulary approach. *PLOS ONE* 2013, **8**:e73791.

39. Zangerle E, Pichl M, Gassler W, Specht G: **#Nowplaying music dataset: extracting listening behavior from twitter**. In *Proceedings of the First International Workshop on Internet-Scale Multimedia Management*. ACM; 2014:21-26.

40. North AC, Hargreaves DJ: **Music and adolescent identity**. *Music Educ Res* 1999, **1**:75-92.

41. Greenberg DM, Rentfrow PJ: **The social psychological underpinnings of musical identities: a study on how personality stereotypes are formed from musical cues**. In *The Oxford Handbook of Musical Identities.* Edited by MacDonald R, Miell D, Hargreaves D. Oxford, UK: Oxford University Press; 2017.

42. Rentfrow PJ, Gosling SD: **Message in a ballad: the role of music preferences in interpersonal perception**. *Psychol Sci* 2006, **17**:236-242.

43. Jacobson K, Murali V, Newett E, Whitman B, Yon R: **Music personalization at spotify**. In *Proceedings of the 10th ACM Conference on Recommender Systems*; *ACM, September: 2016:373*.

44. Zhang B, Kreitz G, Isaksson M, Ubillos J, Urdaneta G, Pouwelse JA, Epema D: **Understanding user behavior in spotify**. In *INFOCOM, 2013 Proceedings IEEE*; *IEEE, April: 2013: 220-224*.

45. Mauch M, MacCallum RM, Levy M, Leroi AM: **The evolution of popular music: USA 1960–2010**. *R Soc Open Sci* 2015, **2**:150081.

46. Russell JA: **A circumplex model of affect**. *J Personal Soc Psychol* 1980, **39**:1161-1178.

47. Watson D, Clark LA, Tellegen A: **Development and validation of brief measures of positive and negative affect: the PANAS scales**. *J Personal Soc Psychol* 1988, **54**:1063.

48. Rentfrow PJ: **The role of music in everyday life: current directions in the social psychology of music**. *Soc Personal Psychol Compass* 2012, **6**:402-416.

49. Gosling SD, Rentfrow PJ, Swann WB: **A very brief measure of the Big-Five personality domains**. *J Res Personal* 2003, **37**:504-528.

50. • Lathia N, Sandstrom GM, Mascolo C, Rentfrow PJ: **Happier people live more active lives: using smartphones to link happiness and physical activity**. *PLOS ONE* 2017, **12**:e0160589.

51. • Warrier V, Grasby KL, Uzefovsky F, Toro R, Smith P, Chakrabarti B, Khadake J, Mawbey-Adamson E, Litterman N, Hottenga JJ, Lubke G: **Genome-wide meta-analysis of cognitive empathy: heritability, and correlates with sex, neuropsychiatric conditions and cognition**. *Mol Psychiatry* 2017.

52. McDermott JH, Schultz AF, Undurraga EA, Godoy RA: **Indifference to dissonance in native Amazonians reveals cultural variation in music perception**. *Nature* 2016, **535**:547-550.