FINAL HONOUR SCHOOL OF EXPERIMENTAL PSYCHOLOGY AND PSYCHOLOGY, PHILOSOPHY AND LINGUISTICS 2020

ENTRY FORM FOR CANDIDATES OFFERING A LIBRARY DISSERTATION

CANDIDATE NUMBER: 1026735

TITLE OF PROJECT:

Harnessing Music Recommendation Systems to Improve Mood Regulation Outcomes in Depressed Individuals

NAME OF SUPERVISOR: Dr Nicholas Yeung

WORD COUNT: 7942

Harnessing Music Recommendation Systems to Improve Mood Regulation Outcomes in Depressed Individuals

1026735

Psychology, Philosophy and Linguistics 2020

Contents

Abstract

Mood regulation is one of the primary motivations reported for personal music listening, leading to numerous attempts over recent decades to exploit associations between sounds and feelings for therapeutic purposes. On the individual level, engaging with sad music in order to achieve catharsis is a common regulatory approach used by those with depression and ruminative cognitive styles. However, recent evidence suggests that this habit does not consistently improve mood, and can, in fact, induce a mood deterioration and an intensification in depressive symptoms. Clinical attempts to improve mood regulation outcomes in depressed individuals through changes to music listening habits have so far achieved mixed results. This dissertation hypothesises that a modern solution for this problem would be to harness the impressive machine learning technology in recommender systems to generate personalised music suggestions that lead to more robust mood improvements. Also posited, is that improving biological plausibility of such a system could further optimise mood regulation outcomes, as distributional semantics methods draw on multiple fundamental aspects of the neural organisation of conceptual knowledge to effectively reflect and guide decision-making. Research has revealed that modelling mood within recommender systems improves their predictive power, and so future investigations could inquire about the effectiveness of mood-based recommender systems on mood regulation in depressed individuals, integrating the most successful principles from traditional music therapies. Collectively, the findings from this dissertation have the potential to advance music-based therapies and facilitate effective use of music listening as a tool for mood regulation.

Introduction

It has been claimed that music exists as a socio-cognitive phenomenon (Koelsch, 2005), transcending cultural boundaries over millennia, but intuitively it remains an immensely personal aspect of our lives that has, for many of us, become inextricable from our memories, aspirations and identities. Operating on the same reward pathways as food, drugs and sexual pleasure (McFerran et al., 2018), music listening is often reported as a form of self-care used to improve one's emotional state (North, Hargreaves & Hargreaves, 2004; Saarikallio & Erkkilä, 2007), but if implemented improperly, the effects may be counterproductive to mood regulation (Garrido & Schubert, 2015b; Garrido, Schubert & Bangert, 2016; Vuoskoski & Eerola, 2012). As will be explored in this dissertation, music has a dual potential with regard to mood outcomes, which is observed most plainly in those who tend towards depression and rumination, who, due to their impaired regulatory abilities (Sakka & Juslin, 2018), are more sensitive to the negative emotion in certain musical themes and features (Garrido & Schubert, 2015b). Engagement with ineffectual coping strategies and a lack of awareness surrounding potential negative mood impacts of sad music means that novel solutions must be proposed that can alleviate rather than exacerbate depressive symptoms. This would be valuable because even momentary distractions from depressive episodes can be of paramount importance to wellbeing (Rutter, 2012). These brief alleviations act to reduce the amount of time spent ruminating and can therefore be preventative of attempts to self-harm (Polanco-Roman et al., 2015), meaning that proper harnessing of music consumption decisions to improve, rather than worsen mood should be of significant concern to mental health professionals.

This dissertation explores some previous attempts to instill healthier music listening habits into vulnerable people, as well as investigating current technologies in music listening and suggesting methods for incorporating them into music therapies. Young people are relying more and more heavily on online content for mood regulation purposes (Bhatia et al., 2019) and personalised content recommendation is becoming ever more integral to the music listening experience. Furthermore, the epidemiology of depression sufferers also skews towards adolescence and young adulthood (Kessler & Bromet, 2013; Watson, 2019). Therefore, this dissertation proposes a depression-orientated recommender system that acknowledges these demographics and recent changes in media use. In order to effectively discuss novel applications for algorithmic music recommendation, it is essential to first provide a background of the semantic representations which underpin their success, as it appears that improving biological plausibility of the system improves its predictive accuracy. Recommender systems attempt to approximate the semantic spaces thought to be part of human neural organisation (Huth et al., 2012). These spaces are where object and actions are represented along certain continuous dimensions according to similarity: such dimensions could include size, colour or animacy. Vast amounts of online corpus data are used to create uniquely specified mappings of the semantic space onto relevant judgement criteria. By adequately modelling music-mood, music-person and person-mood relationships into a recommender system, it could be possible to improve low moods in the short term, or even reduce pathological symptoms of at-risk populations through mood-enhancing personalised music recommendations.

The Nature and Impact of Music Listening Choices

Determinants of Music Preference

During the course of an average day, it is common that one will make multiple conscious or unconscious evaluations about music, and it is important to understand the cognitive and contextual variables that inform those judgements in order to make claims about how such factors could be exploited to achieve mental health objectives. Some of the broad and stable determinants of music choice are already known and have been empirically evidenced across cultures. One immediately plausible candidate is that of personality trait dimensions (Kim et al., 2019). For instance, when contrasting purpose of use between introverted and extraverted individuals, extraverts tend to use music as a tool to encourage socialisation (Schäfer & Melhorn, 2017). Other studies that have endeavoured to tackle the question of whether the Big Five personality dimensions are related to music preference have varied hugely in their methodology and their success at establishing a link between the two, measuring preferences for different genres, rhythms, tempos or feelings. A meta-analysis of these attempts covered 28 studies and over 250 000 participants, and found weak, mostly non-significant relationships between personality dimensions and style preferences with an average correlation of only $|r| = 0.058$ (Schäfer & Melhorn, 2017). The strongest correlation was for the association between openness to experience and sophisticated music (encompassing blues, jazz, classical, etc.) at $r = 0.21$, meaning that the factor only explained around 5% of the variance in preference for that class of music (Schäfer & Melhorn, 2017): all other correlations were notably smaller.

However, many of the studies included categorised music using to a genre taxonomy which is notoriously ill-defined, suffering from unclear category boundaries, low ecological validity and high degree of subjectivity (Brisson & Bianchi, 2019). Genres can include a huge stylistic range, for example with “oldies” being grouped solely by age and “soundtracks” often having very few shared features outside of their usage, suggesting that this type of classification may be unhelpful for a rigorous scientific investigation of preference. A potential first step in elucidating all of this contradictory and often poor-quality data, would be to perform another meta-analysis disregarding all of the genre-based classifications to observe if any of the other music attribute models could uncover personality dimensions as a consistent determinant of taste. Other classifications have produced far more consistent evidence for demographic variables as preference factors, such as the MUSIC model that identifies five unique preference dimensions – mellow, unpretentious, sophisticated, intense and contemporary (Schäfer & Melhorn, 2017). Age, gender and culturally informed self-views have all demonstrated significant impacts on preferences for these classes of music (Bonneville-Roussy et al., 2013; McCown et al., 1997; Sumare & Bhalke, 2015). For example, self-view is an essential factor in determining which forms of expression are deemed most critical in reaching a sense of individuality; one might be attracted to alternative or avant-garde music if they view themselves as unconventional, or energetic music if they perceive themselves as athletic (Rentfrow & Gosling, 2003). Though much of this evidence is compelling and indicates that relationships exist between stable listener factors and their music preferences, a simple way to improve the methodology here could be to map out music preferences in a continuous semantic space and therefore bypass the need to generate potentially problematic categories. In doing so, preferences can be asserted for regions within that semantic space that may cross boundaries between traditional genre or MUSIC categories.

Interactions Between Music and Mood

General Interactions

Although many of the major stable preference factors have so far been identified, there are a number of more transient, context-based elements yet to consider. The most relevant of these contextual factors to our discussion here is mood – a short-lived emotional state that lacks an object of arousal ("APA Dictionary of Psychology", 2020). This differentiates moods from emotions which are elicited by a particular stimulus in a more pronounced fashion, such as a threat arousing fear. Music and mood have a long-established and multifaceted relationship where mood can impact music listening decisions and music can express and induce emotional states in a listener (Dolan & Sharot, 2011): ideas which have been harnessed in some successful music therapies (Aalbers et al., 2017; Erkkilä et al., 2011; Garrido & Schubert, 2015a). Musical elements like tempo, rhythm, timbre and harmony may be used to either reflect or counteract one's mood (Sumare & Bhalke, 2015); for example, an individual may choose slower, more harmonic music to relax when feeling anxious, or may choose faster, more upbeat music to maintain a positive mood. There are also spectral features of the sound wave that we are less consciously aware of, but which still impact mood outcomes, with higher spectral flatness, greater noise and more spectral flux all leading to mood deterioration (Han et al., 2009).

Sad Music for Mood Regulation

One interaction of particular consequence and complexity is that of sad music and mood – the intersection of which is portrayed in starkly different terms by the scientific community and stories in mainstream media. Mood regulation, specifically entailing mood improvement and control, has been evidenced by quantitative and qualitative findings as a primary motive behind listening to music, and sad music in particular (North, Hargreaves & Hargreaves, 2004; Saarikallio & Erkkilä, 2007). The hypothesised benefits of listening to sad music, containing negative lyrical themes and/or acoustic cues associated with sadness, have been talked about since the days of Aristotle, who suggested it could be overwhelming to a point of catharsis (Kawakami, 2013). Since then, those ideas have propagated to a point where societies broadly accept the curative properties of wallowing in one's own sadness through music and portray the idea unchallenged in mainstream film and television. Such a pattern is observed cross-culturally, with listeners drawing on mostly universal cues for sadness in music like slower tempos and dissonance (Balkwill & Thompson, 1999; Swaminathan & Schellenberg, 2015; Taruffi & Koelsch, 2014)

However, recent data suggests that following these media tropes too stringently can be counterproductive, leading to a deterioration in listener mood. In most other forms of media use, there is consensus in harnessing uplifting content to improve one's mood, with premenstrual women consuming more comedy programmes and positive news stories as a form of mood management (Biswas, Riffe, & Zillmann, 1994; Meadowcroft & Zillmann, 1987): this appears functionally opposite to the use of sad music for purposes of mood enhancement. Though there are plenty of reports that show sad music is a legitimate and useful tool for cognitively processing negative emotions (Garrido, 2017) and safely exploring deeper feelings, the results for mood impacts specifically have been significantly less consistent. Along with numerous studies that have shown happy music to work better than sad music as a mood enhancing tool (Garrido, Schubert & Bangert, 2016; Vuoskoski & Eerola, 2012), there are also some reports that claim little to no negative feeling from sad music (Manuel, 2005), which may be a result of methodological drawbacks. Some of these findings are difficult to compare due to unclear or undiscerning usage of the terms “mood” and “emotion”, resulting in conclusions which cannot adequately be assigned to either measure. The measures themselves are also problematic, as they consistently involve explicit measures of mood such as self-report questionnaires which are subject to motivational biases (Hofmann et al., 2005). Investigating these stimuli using implicit alternatives such as the IPANAT or the IAT could provide a better understanding of the unconscious impacts on mood from listening to sad music.

Mental Health Outcomes

Sufferers of mood dysregulation symptoms are the group that receive the strongest and most consistently negative impacts on their mood from listening to sad music, including individuals with depression and those with a ruminative cognitive style (Garrido & Schubert, 2015b). Rumination is the tendency to focus negatively on the events in one’s life and contemplate the potential causes and consequences therein. The trait is highly predictive of depression, which is characterised by persistently low moods and a diminished interest in typical activities (Truschel, 2019). Those with depression or a ruminative cognitive style, for which there is significant crossover, often engage in maladaptive processing of negative affect states that tends to prolong them in length and hinder more effectual coping mechanisms (Garrido & Schubert, 2015b). Negative attentional biases impede the normal detachment from negative stimuli, leading to findings that ruminators interpret neutral facial expressions more negatively than controls (Raes et al., 2006) and show disproportionately high mood-congruent recall of negative words (Matt, Vázquez & Campbell, 1992). Such biases, in conjunction with lower motivation to engage in mood enhancing behaviours (Garrido & Schubert, 2015b), facilitate a propensity to interpret music, even emotionally neutral music, more negatively (Hunter, Schellenberg & Griffith, 2011) whilst also engaging with this mood-congruent music more than the neurotypical population (ter Bogt et al., 2019).This heightened sensitivity to, and overconsumption of, negativity in music has resulted in many parallel findings that sad music worsens feelings of dysphoria and social isolation in people with unhealthy thinking patterns (ter Bogt et al., 2019; Garrido & Schubert, 2015a; Garrido & Schubert, 2015b; Rentfrow & Gosling, 2003), an effect which is modified by duration of listening and intensity of the initial mood (Garrido, Schubert & Bangert, 2016).

Though little investigation has been done concerning the long-term mood and mental health impacts of listening to happier music, there is evidence to suggest that ruminators are also atypically sensitive to the short-term mood enhancing effects of happy music (Garrido & Schubert, 2015a). When a sample of adolescents accessing mental health care were guided by music therapists to be more conscious in their approach to music listening, they showed a decrease in psychological distress, reporting improved feelings of personal agency and hopefulness (McFerran et al., 2018). The most succesful and long-lasting mood improvements in music therapies appear to follow a tenet known as the iso-principle (Davis, Gfeller & Thaut, 2008; Heiderscheit & Mason, 2015) – a process in which music selections begin as mood-congruent (i.e. negative) and gradually become more positive over the listening duration by controlling elements such as the pitch and tempo (Clark & Tamplin, 2016). However, when music listening habits become more rigidly enforced, in the form of “prescribed playlists” and the like, the benefits to mood become more limited (Garrido et al., 2016).

In light of this evidence, and the consideration that more conscious music consumption may be difficult for those who tend towards depression, the optimal strategy may be to provide a framework which bypasses the need for human supervision and provides music suggestions in a manner that follows the iso-principle. Fortunately, such a system could now be viable by assimilating these principles into the increasingly high quality music recommendations that streaming services can personalise using distributional semantics. In order to propose a system for mood regulation, it is important to first consider how music, and conceptual knowledge more broadly, is organised within the brain and how representations might be impacted by a contextual factor like mood, as these are the properties that recommenders are attempting to replicate. Indeed, as will be explored, a more biologically plausible system would appear to produce better personalised recommendations.

Conceptual Knowledge as a Continuous Semantic Space

Semantic Memory in the Brain

Organisational Principles

While there is relatively little literature regarding the specific representation of music in the brain, there is substantially more on the domain general principles that apply to semantic information across all classes which are important to represent in computational models like recommenders. The behavioural, neuroimaging and computational evidence points towards a distribution of semantic knowledge in a continuous space, the understanding of which will be key to navigating from more negative to positive affective states. Early theories about the potential representations of conceptual knowledge were based around the notion of a semantic hierarchy which stored information from most general to most specific (Rogers & McClelland, 2004): e.g. from music to rock music to The Rolling Stones. Developmental data has suggested that general information about concepts was acquired first and the more fine-grained category differences were not completely understood until later on at age 11 (Keil, 1979). Correspondingly, patients with semantic dementia who suffer from generalised deficits display an opposite pattern of degeneration, losing their grasp of subordinate level distinctions first (Warrington, 1975).

Alternatives that better consider typicality and flexibility in concepts have since proved more suitable, indicating that category membership is continuous as opposed to discrete. Research shifted focus onto findings that some objects and actions are rated as more prototypical category members than others (Rosch, 1973) and that there are psychological effects induced by the level of family resemblance. More prototypical members are verified as part of their category more rapidly (Rips, Shoben & Smith, 1973), they are acquired earlier (Mervis, 1987) and tend to be lost later in semantic dementia process (Warrington, 1975) which suggests this gradation in category membership. Neuroimaging has consolidated this notion of continuous semantic dimensions, with significantly graded activation occurring in the ventral vision pathway in response to 10 different animal stimuli that varied in their animacy. The smaller inter-stimulus differences revealed a non-dichotomous relationship, with the least animate stimuli closely mirroring responses to artefacts, and a steady decrease in response similarity as stimuli became more animate (Sha et al., 2015).

Concept representations are also less fixed in the sense that there is flexibility in their contextualised manifestation – whilst there are often concrete elements to concepts, the specifics of the mental image conjured up may differ depending upon sentence or interpreter variables. Semantic priming is a clear consequence of this flexibility, allowing for homonymic or polysemic objects to be identified faster than they would without any kind of contextual cues. The representation of a word such as “man”, for instance, must be sufficiently flexible to account for references to the entirety of the human species, the male of the human species and the adult male in particular. Presented in a vacuum, the precise meaning is always unclear, and so prior activation is integral to successful comprehension of near-infinite potential exemplars (Solomon, Medaglia & Thompson-Schill, 2019).

This extends to contextual understanding of similarity relations, as demonstrated behaviourally by Macario (1991) who presented unfamiliar objects to infants and were then primed with knowledge about their basic level category. Reasoning with the same concept was dependent upon whether the object was primed as a food or as a toy – there was domain-specific attribute weighting in that food was generalised based on colour and toys were generalised based on shape (Macario, 1991). Shifts in representation can also be observed more directly in the patterns of neural activity themselves, varying as early the first sensory inputs to the V4 and later, during cognitive processing, on the basis of attentional changes (Çukur, et al., 2013). This expansion of relevant representations in favour of less relevant ones is a process displayed by depressed individuals in the form of negative attentional biases where negative stimuli are expanded over neutral stimuli (Platt, Murphy & Lau, 2015). In the context of sad music, this could perhaps insinuate that the representations of negatively affecting musical elements and themes are atypically warped at the cost of more positive sounds: investigation using neuroimaging techniques such as fMRI would be useful to test such a hypothesis. Overall, this evidence reveals some of the properties necessary to create a biologically plausible computational model of semantic memory, with continuously graded representations that can be influenced by contextual factors.

Neurobiology

In terms of neural organisation, it is broadly accepted that semantic knowledge has some underlying structure that is common between individuals, with similar concepts eliciting similar patterns of activity (Thompson-Schill et al., 1999). A continuous semantic space would allow for category similarity to be represented by physical closeness within the cortex (Huth et al., 2012), meaning that, for instance, R&B musical features would be represented in closer proximity to soul than to classical music elements given the relative conceptual distances. Using principle components analysis of fMRI data which was matched with language data from an online lexicon called WordNet, evidence for a shared semantic space (see Figure 1) was found with graded regions specific for representing animals, humans, actions, outdoor spaces and motions (Huth et al., 2012). This aligns with data from supposedly category-specific brain regions like the extrastriate visual cortex (commonly known as the fusiform face area) which has been shown to respond significantly to perceptually similar non-face stimuli (Gauthier, Berhmann & Tarr, 1999). Therefore, the neurobiological evidence consolidates that notion that boundaries between category representations appear, like the categories themselves, to be somewhat porous and that semantic knowledge is likely represented continuously in the brain.

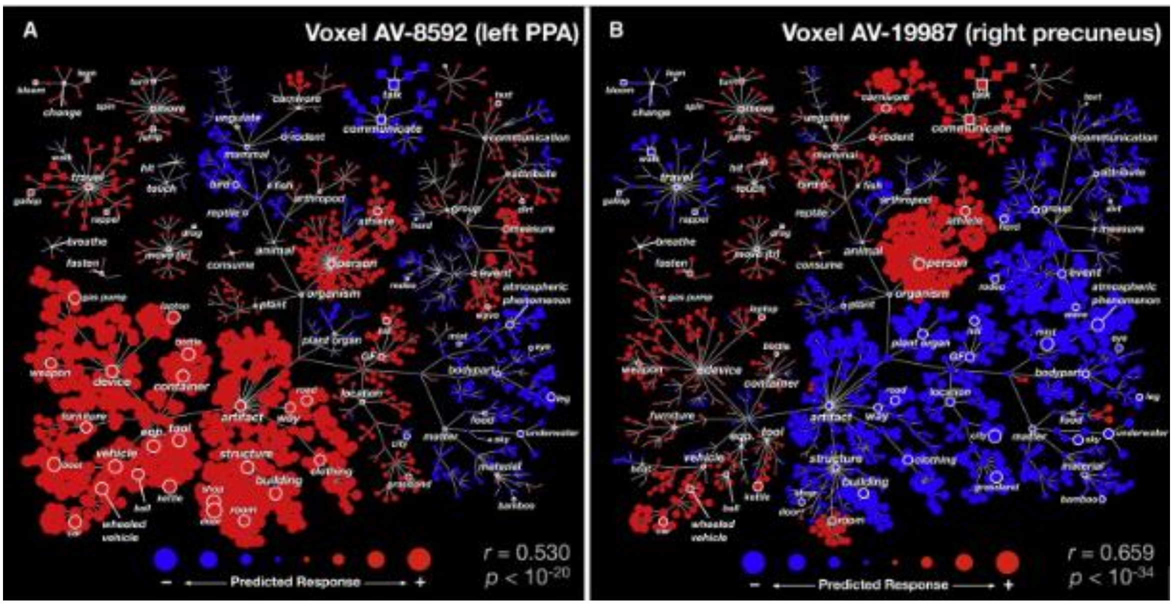

Figure 1. The predicted responses for two individual voxels for each of 1705 categories with links between nouns (circles) and verbs (squares) indicating a relationship and marker size indicating predicted response magnitude. Prediction accuracy is given as (*r)* – (A) appears to be selective for man-made objects and (B) is selective for people and animals. Reprinted from A continuous semantic space describes the representation of thousands of object and action categories across the human brain, by A. Huth, S. Nishimoto, A. T. Vu & J. L. Gallant, 2019, *Neuron, 76(6)*, p. 20.

Domain-General Representations

Most concepts have multimodal aspects to how they must be represented and understood – knowledge of a musical instrument includes sensory information about what it looks like, how it sounds and how it feels to play. These different elements of a concept are represented in activity specific to various neocortical regions that process perception and action, but they are thought to be mediated by a supramodal hub which abstracts away from modality-specific attributes (Mollo et al., 2017; Patterson, Nestor & Rogers, 2007). According to this "distributed-plus-hub" view of semantic memory (see Figure 2), the hub is situated in the anterior temporal lobes and links all of the modality-specific regions, or spokes, via bidirectional white matter connections (Ralph et al., 2017). The notion of a semantic hub that transcends modality arose from findings surroundings semantic dementia, where it was discovered that patients displayed semantic impairments across all task types, concept types and modalities with the exception of simple numerical understanding (Ralph et al., 2017). The atrophy to the anterior temporal regions had, therefore, caused the deterioration of concepts in their entirety and not just how they look, or sound, or smell. Domain generality in ATL representations is further consolidated by investigations of its correct functioning – neuroimaging findings indicated that along with specific activation of the left inferior parietal lobule for word stimuli and the right middle occipital gyrus for picture stimuli, there was common activation in the temporal regions (Vandenberghe et al., 1996).

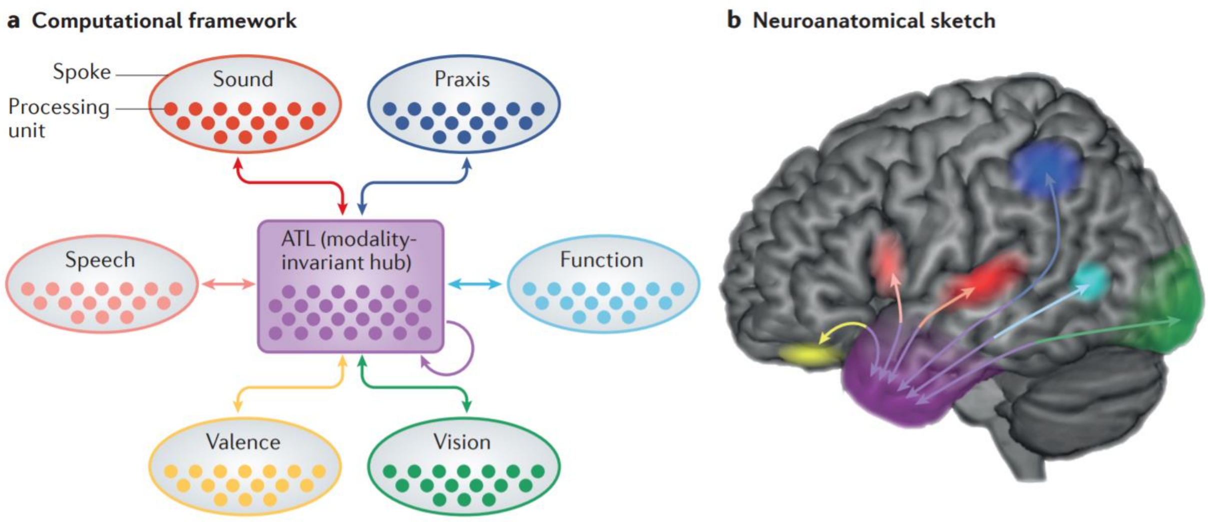

Figure 2. The computational (A) and neuroanatomical architecture of the hub-and-spoke model. Reprinted from The neural and computational bases of semantic cognition, by M. Ralph, E. Jefferies, K. Patterson, & T. Rogers, 2017, *Nature Reviews Neuroscience*, *18*(1), p 2.

Music-Specific Representations

The direct evidence surrounding how music is represented in the brain shows that various elements of the distributed-plus-hub model may be harnessed for such processing. There is, of course, no true functionally localised music centre, only an array of components from the general auditory network that are often harnessed for language processing but can interpret musical stimuli (see Figure 3). Extraction of musical features like dynamic pitch variation, timbre and intensity occurs primarily in the right subregion of the ventrolateral superior temporal gyrus according to neuroimaging and lesion studies (Jentschke, 2007; Johnsrude, Penhune & Zatorre, 2000; Scott et al., 2000). This area interacts with working memory in the frontal and temporal cortices, as well as the premotor cortex to produce online representations of the tonal and metrical structure of musical material (Zatorre & Salimpoor, 2013). As assumed by the distributed-plus-hub model, there is a trajectory of auditory information that projects from the relevant primary sensory regions to the anterior temporal region where the N400 activity was shown to converge during a multimodal semantic judgement task (Marinkovic et al., 2003). Ultimately, therefore, the structure of music representations appears to align with semantic knowledge from across other domains.

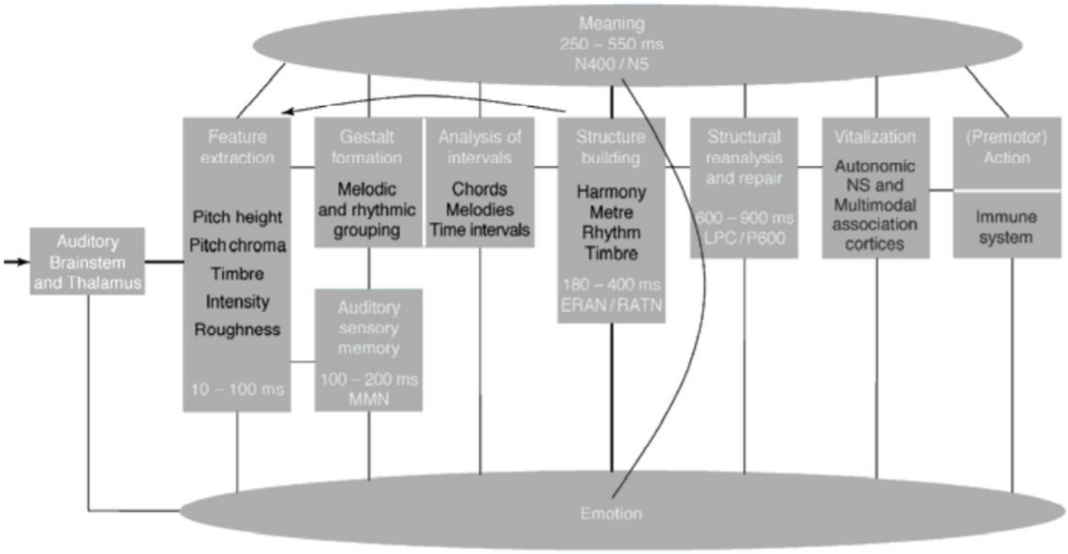

Figure 3. A neurocognitive model of music perception. Reprinted from *Neural Correlates of Processing Syntax in Music and Language: Influences of Development, Musical Training, and Language Impairment* (p. 16), by S. Jentschke, 2007, Unpublished Doctoral Thesis.

Computational Models of Semantic Memory

Computational models of this continuous space can now adequately account for many of the fundamental properties that have been observed in acquisition, deterioration and operation of semantic knowledge (Rogers & McClelland, 2004). In Rumelhart's initial connectionist formulation, the model had inputs of items and relations, and outputs of item attributes (Rumelhart, 1990). Models like this are composed of many interconnected units, like neurons in the brain, with concepts being represented as unique patterns of activity that are distributed across multiple units within the network. To begin with, the weighted connections between units are random, but develop via backpropagation in a gradual, iterative process that is demonstrated by a decrease in error and the pathways between becoming stronger or weaker as co-occurrence dictates.

Distributional semantic models differ slightly in that they learn meaning through episodic experience in a linguistic environment such as an online corpus (Jones, Willits & Dennis, 2015). This allows the model to accumulate statistical information regarding how word usage is distributed, representing a lexicon according to Harris's distributional hypothesis (1970) that posits contextual overlap of similar terms. As with connectionist models, distributional models generate latent space mappings where distance reflects conceptual similarity: two possible dimensions could, for instance, be age and ethnicity, grouping people of the same demographics together in space. Though research has converged on this notion of a representational space, there is not yet a complete correspondence with biological representations, since models rely so heavily on raw statistics, overlooking important facets of meaning like causal relations and sensorimotor properties. In spite of this, along with somewhat implausible learning data, computational modelling of semantic knowledge is sufficient that such a system can be incorporated into algorithmic content recommendation systems. Though connectionist networks are central for understanding how individuals might learn concepts, distributional networks are ultimately much more powerful and have far greater predictive accuracy for decision-making. Hence, they are preferred in recommendation systems and could provide an effective and somewhat biologically plausible platform for mood regulating applications.

Music Recommendation Systems and their Mental Health Applications

Modern Recommender Systems

Purpose

When consuming media, people will regularly, though sometimes unconsciously, engage in selective exposure to online content for mood regulatory purposes, seeking refuge in a virtual world with alternative affective experiences (Schramm & Cohen, 2017). However, individuals often arrive on entertainment platforms with no articulated desire for their media consumption (i.e. no direct search queries), which is shown in the fact that 60% of YouTube clicks from the homepage are onto recommended videos (Davidson et al., 2010) and 30% of Spotify listening is recommended tracks ("HBS Digital Initiative", 2018). Media recommendation has seen a significant shift since the advent of computers and digital content, away from traditional sources like critics, and towards a personalised, algorithmic approach. Personalising media consumption to this degree began primarily as a response to the vast overload of online content that continues to grow every day on the myriad of streaming, journalistic and commercial platforms that are now available to us, and continues to continues to garner scientific and economic interest. This has required the development of ever more efficient pathways to accurate preference prediction and relevant content discovery that can effectively facilitate the mood regulation that is desired (Aucouturier & Pachet, 2003).

Computing Similarities

Collaborative Filtering.

The consistency with which people turn to personalised recommendations to inform their online decision-making has grown in response to the increasing efficacy of the machine learning used – a type of learning that often utilises a technique known as collaborative filtering (Shardanand & Maes, 1995). This technique was one of the first developed for algorithmic recommendation and has since become one of the most prevalent in commercial use, but recent investigations suggest that incorporating latent semantic mapping into the system can greatly improve its predictive accuracy. Despite the nomenclature, there is no true interaction between users in these systems, it is only collaborative in the sense that the system is analysing individuals' preference data and using them to compute affinities between users (Alaimo & Kallinikos, 2019). Modern consumers of entertainment have become critics in their own right as a result of the numerous feedback options the internet has provided us; there are websites like IMDB and Rotten Tomatoes that are purpose-built for explicit numerical feedback with their own rating systems, and there are options to "like" or "dislike" content on all the mainstream streaming services. Further, music companies like Spotify can incorporate implicit information such as streaming numbers, additions to user playlists and visits to an artist's discography following exposure to their music (Ciocca, 2017). Collaborative filtering takes data like this to measure levels of similarity in preference between individuals and suggest content based on what an individual has enjoyed, but what another similar individual has not yet come across. The process, therefore, relies on evidence that tastes are homogenously distributed in general (Blattner, Zhang & Maslov, 2007), assuming a stability of preferences over time and disadvantaging individuals with more diversity in their preferences (Alaimo & Kallinikos, 2019).

Within a music recommender such as Spotify's Discover Weekly system, the preference data is organised into a matrix with rows of users and columns of tracks which is then factorised (Blattner, Zhang & Maslov, 2007). The resulting vectors can be compared to one another to determine the degrees of similarity between users and between tracks (Ciocca, 2017). Researchers have repeatedly compared the item-based against the user-based approaches and have consistently concluded that, while notably less efficient, an ensemble approach is significantly more accurate in its predictions than either of the other two approaches are by themselves (Azzopardi, 2018). Bringing in other methods at this stage is particularly important, given that collaborative filtering is so limited by what other listeners have listened to: semantic mapping can account for more fundamental and context-sensitive similarity.

Raw Audio Analysis.

Another technique often used in combination with collaborative filtering to form hybrid recommenders analyses the raw audio data itself and produces recommendations on the basis of similarities between the musical elements (“HBS Digital Initiative”, 2018). By suggesting music in this manner that does not require user ratings or discussion, the service is now capable of recommending tracks that have only a dozen listens alongside those that have millions of plays, thus creating a slightly fairer creative environment in which smaller and newer artists can flourish. The convolutional neural network uses technology comparable to that of facial recognition, but adjusted for audio stimuli (Ciocca, 2017), to scan through the spectrogram and extract the statistical data that corresponds with characteristics like tempo, loudness, key and time signature (Ciocca, 2017; “HBS Digital Initiative”, 2018). The acoustic analysis allows for each event in an individual track to be understood (each chord, each crescendo, each solo, and so on) and locates other tracks that contain similar events or event progressions (Prey, 2018). Aggregating all of these features allows for automatic groupings to be formed which parallel semantic spaces in the human brain, extracting the latent factors and producing subcategories for different rhythms, modes and timbres.

Latent Semantic Analysis.

The third pillar of modern music recommendation involves natural language processing which incorporates latent semantic analysis (LSA) as a tool, therefore making direct use of semantic spaces to predict preferences. When basing predictions on user-item ratings, the result is essentially an approximation of the underlying semantic connections between items (Ruiz-Montiel & Aldana-Montes, 2009). This why directly measuring and modelling distributional semantics of the tracks is becoming ever more common in commercial recommender systems. Streaming companies like Spotify have augmented their recommender systems by collating written data from all over the internet to provide a language-based representation of each track based upon the keywords that are used in reference to them (Ciocca, 2017). The most frequently used terms are then weighted and used to locate other artists or tracks that are defined similarly, meaning that their respective keywords occupy adjacent areas in the semantic space. Therefore, as in the brain, these entities can be represented as the overall activity of their elements in relation to subspace dimensions. Where the audio analysis may have found more likeness between Elton John's "Candle in the Wind" and Simon & Garfunkel's "Bridge Over Troubled Water", with their acoustic pianos and mellow vocals, the LSA approach could define "Candle in the Wind" in relation to Princess Diana or the charity work that it funded, potentially pulling it closer to a song like Band Aid's "Feed the World". Therefore, recommendations can be made from the relevant subspaces of the music catalogue as opposed to relying on the traditional, but problematic and ambiguously defined genre pools which contain items that may have very few commonalities.

A somewhat problematic assumption of LSA in this context, however, is that of representational homogeneity. In amalgamating language data from a vast sample into a single network and using that distribution to predict decision-making, it is presumed that all participants will have the exact same knowledge and will have had the exact same experiences with all of the entities included: i.e. that their representations are the same. This said, it would be unfeasible to create distributions based only on individual language data, as the coverage of concepts would be far sparser, and in creating this averaged semantic memory, individuals with less experience in a choice domain can still receive accurate predictions. There are no currently published figures for the predictability of these three recommendation models working together, but it appears highly likely that they have provided a considerable improvement over the collaborative filtering technique alone, given the popularity of the services which have combined them. It would be of interest for researchers to investigate this issue directly, systematically varying the relative weighting of each individual model to find the appropriate balance for optimising the predictability of the music recommendations outputted.

Generating Predictions

Recently introduced techniques have suggested that the continuous spaces arising from distributional semantics can be used to generate predictions of multiattribute choice, thus providing methods to operationalise the affinities that result from latent semantic analysis and raw audio analysis. In 2019, a group of researchers attempted to predict association-based judgements across multiple choice domains using a technique called vector mapping which, as shown in Figure 4, uses training data on a distributional semantic network to locate the regions in the semantic space which are most associated with the judgement dimension. The predictive accuracy of vector mapping was compared to a baseline approach of vector similarity, where only relative similarity of an object to words signifying the high and low ends of a judgement dimension is considered. Judgements were made concerning the relative warmth of famous people, nutritional values of food, utility of certain goods, and so on, ensuring breadth of analysis (Bhatia et al., 2019). Vector mapping proved to be an impressively powerful predictive tool, with an average correlation of 0.77 between predicted and actual judgements and statistically significant results across all behavioural domains: this far outperformed the vector similarity approach as well as traditional multiattribute decision rules (Bhatia & Stewart, 2018; Bhatia et al., 2019). Recommender systems can therefore represent what people already know about entities and their attributes with latent semantic analysis and raw audio analysis, along with how that knowledge would be used to decide whether or not they will consume certain content using a technique like vector mapping.

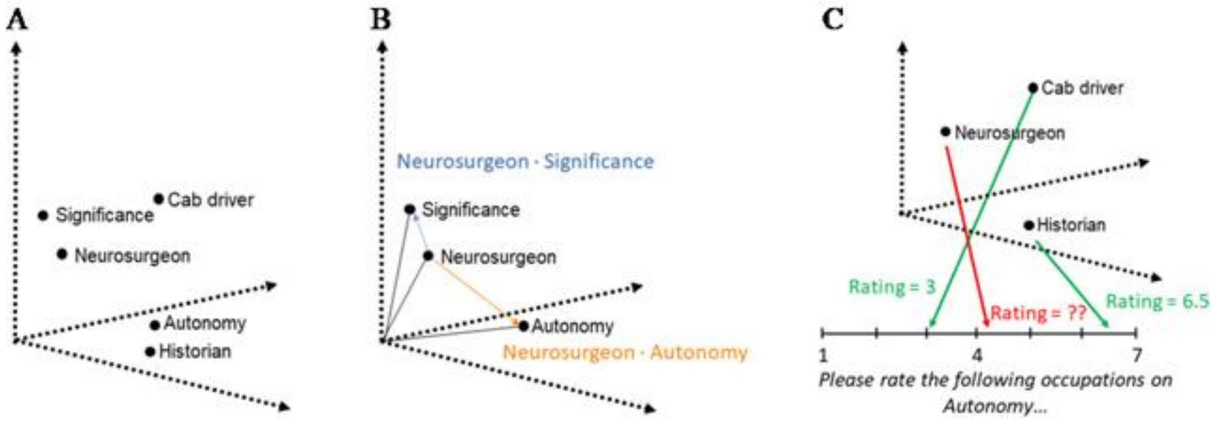

Figure 4. A visualised comparison of vector similarity (B) and vector mapping (C) approaches to predicting judgements about occupations from three-dimensional semantic space (A) – vector mapping includes a separate scale for the relevant rating dimension. Reprinted from Predicting High-Level Human Judgment Across Diverse Behavioral Domains, by S. Bhatia, R. Richie, W. Zou & S. Vazire, 2019, *Collabra: Psychology*, 5(1): 50, p. 3.

Modelling Contextual Factors in Recommender Systems

Environmental Factors

Until the recent “contextual turn” (Pagano et al., 2016), recommender systems would assume user preferences to remain static and immutable regardless of any situational factors, disregarding the notion of situated actions and losing predictive power as a result (Han et al., 2009; Prey, 2018). By modelling contextual factors such as time and location, the system is better optimised to recreate the behaviour of biological cells which is thought to be valuable since machine learning tends to perform better when the system resembles the neurons it is attempting to emulate. In an investigation of the visual system, it was discovered that within a class of biologically plausible neural networks, there was a strong positive correlation between success at image categorisation and the model’s ability to predict neural responses in V4 and the inferior temporal cortex (Yamins et al., 2014). Biological plausibility in this case, entails capturing the impact of contextual factors on the semantic mapping of tracks, adjusting the outputted predictions to account for the environment. In doing so, the system could, for instance, accurately recommend energetic, high tempo music to an athlete training at the gym, but otherwise present them mellow and acoustic sounds if that were their standard preference.

Context awareness begins at the data compilation stage, which is now easier than ever before, considering that a large proportion of media consumption occurs on mobile devices which regularly track information such as time, GPS location, weather, level of activity, and even traffic conditions (Adomavicius et al., 2011; Prey, 2018). Of course, obtaining the data implicitly is not flawless, leaning heavily on inference and sometimes leaving questions of motivation unanswered, but the sheer quantity of data is exceedingly useful for the variables that do not require such parsing. Once the contextual data is collected, it must be integrated into the system in order for two-dimensional rating functions to become three-dimensional (i.e. user-item-context), thus producing ratings that can be used for more flexible recommendations. Contextual modelling in the music recommendation domain has resulted in predictive accuracy increases of between 7 and 33% over traditional methods (Baltrunas et al., 2011; Prey, 2018). This demonstrates that biological plausibility is a key property of well-designed machine learning across cognitive domains, and that more accurate replication of biological semantic memory through context sensitivity can improve recommender systems.

Mood

Another major contextual factor that requires modelling within recommender systems is that of user mood, which in turn necessitates an adequate method for classifying moods in music and identifying moods from user data. From the starting point of a music streaming service that performs latent semantic analysis and raw audio analysis on its item catalogue, the system must vector map each item’s features onto a particular mood dimension. Multiple models of mood exist, but the dimensions used by researchers have often been based upon Thayer’s model, presented in Figure 5, which represents emotional states in a two-dimensional plane with arousal and valence axes. The plane is subsequently divided into eleven mood categories: angry, excited, happy, nervous, calm, pleased, bored, relaxed, sad, sleepy and peaceful (Baltrunas et al., 2011). The acoustic features that are extracted from the primary analysis are then translated onto one of those categories using a technique like support vector machines (Seo & Huh, 2019). This technique involves a survey matching pieces of music with the emotional value perceived by the listener is completed and a linear regression is performed to establish the relationships between musical elements and moods: intensity, tempo, rhythm, etc. are the predictor variables for response variables of arousal and valence (Seo & Huh, 2019). Therefore, when a new track is inputted into the system, its features are automatically extracted, and their values are used to map onto the relevant category in the mood space and provide a classification estimate. This process can be repeated using the latent semantic space to generate a second classification that can make use of important information that is not captured in the raw audio alone: elements such as the lyrical theming or the context surrounding an album’s release. Combining both types of vector mapping is, therefore, a practical method of generating a fuller representation of a track’s mood value.

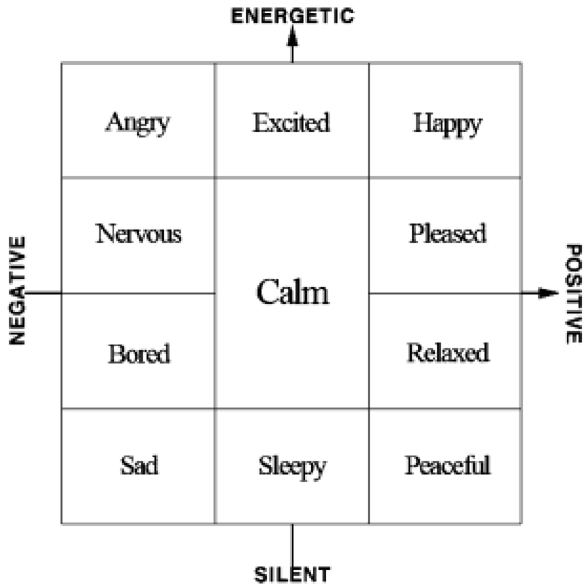

Figure 5. Thayer’s two-dimensional representation of mood with arousal on the y-axis and valence on the x-axis. Reprinted from Music emotion classification and context-based music recommendation, by B. Han, S. Rho, S. Jun, E. Hwang, 2007, *Multimedia Tools And Applications*, 47(3), p. 435.

A version of this, using only audio analysis, was incorporated as part of a context-aware music recommender which had the explicit goal of transitioning the user to their desired emotional state (Han et al., 2009). After the features were extracted and the matrix was factorised to reduce dimensionality, participants completed an online questionnaire regarding their shifts in mood while listening to musical stimuli. A support vector machine algorithm was then applied to relate music feature information to the emotion transition data, creating a model that could learn to predict music with mood enhancing features with an overall accuracy here of 68% (Han et al., 2009). Accuracy here refers to how many correct prediction outputs the classifier can predict across all the data. This study, among numerous others, has concluded that modelling mood consistently improves recommender system precision (Andjelkovic, Parra & O'Donovan, 2016; Meng et al., 2018; Wang et al., 2018),

While certainly an improvement upon early recommender performance, there is still a significant discrepancy between predicted and actual preferences that remains unaccounted for. There are a number of elements within these systems that could be further optimised or reconsidered to more accurately reflect the interactions between people and the music they listen to. Firstly, most systems currently assume a uniformity in emotional value over the duration of a single track – the significant fluctuations in mood during songs such as Queen's "Bohemian Rhapsody" cannot not adequately be captured by feature analysis, thus complicating response prediction. Secondly, and most importantly, the methods used to collect data on user mood and item mood are overly cumbersome and imprecise. The use of self-report measures to ascertain the emotional value of musical features entails a certain level of motivational bias that is informed by ever-changing and culturally specified knowledge surrounding sounds and their meanings. Self-report is also ineffective for investigating listener mood, in that it is often difficult to define one's own mood using language, and there is often a complexity that cannot be captured in a single word.

The use of smartphone sensors to determine context appears to be the next significant upgrade to recommender precision, showing promise in other domains such as tourism (Meehan et al., 2013) where location tracking has been harnessed. Inference via environmental variables has been an important first step, but with recent advances in machine learning, mood can now be identified more directly by algorithms that pick up on facial micro-expressions or speech features that are indicative of mood (Erdal, Kächele &

Schwenker, 2016; Grimm, Kroschel & Narayanan, 2007; Li et al., 2017; Pfister et al., 2011). In future, these techniques could be employed by music recommenders by incorporating such technology into smartphone cameras and digital assistants, enabling continuous and individualised data collection for both implicit user mood and emotional response to musical cues while the user listens to their music. There have already been reports of companies like Huawei and Apple developing assistants that are capable of mood recognition (Burlacu, 2018; Davenport, 2018), so the possibility of integrating this data for music recommendation by streaming services appears to be on the horizon. Therefore, it is possible for the specific contextual factor of mood could be represented within a recommender to create a more biologically plausible and predictively accurate system.

Applications to Depression

To borrow a metaphor from Tim Quirk, Head of Global Content Programming at Google, the role of music platform designers is now that of park ranger (Reskinoff, 2014), who must tend to a vast musical landscape and ensure their visitors enjoy their excursions by promoting healthier interactions with the elements of that landscape. As the flora and fauna continue to grow in quantity, the park rangers must provide some way of navigating the ever-expanding terrain, not by suggesting their own, personal favourite paths, but by providing options that will allow the visitor to maximise the benefits they gain from their experience in the park. In practice, this guiding role is of greatest importance to the more vulnerable among us – depressed individuals, and those who are at a higher risk of diagnosis, are significantly more sensitive to negative emotional values in music, leading to mood dysregulation and the potential for exacerbation of pathological symptoms in the long-term. Therefore, using the most advanced technologies available, the powerful emotional force of music could be harnessed to steer these listeners towards music habits that promote recovery from depressive disorders, rather than passively facilitating damaging behaviours by only recommending more sad songs.

From a starting point of standard music therapy techniques that use the iso-principle, it is possible to build on this fundamental notion and assimilate it into a form of individualised mood regulating content provision that could be offered on a vast scale. The tools are available for major music streaming services to accurately model the mood of a listener and the mood of a song, and frame these measures as part of their recommender system to provide highly accurate mood-based track suggestions. These recommendations could be operationalised in accordance with the iso-principle, slowly shifting from negative to positive emotional values, thus providing a gentle push towards mood enhancing behaviour that will provide longer sustaining improvements that still allow for enjoyment and cognitive benefit to be obtained from sad music initially. A music streaming platform could combine traditional collaborative filtering, raw audio analysis and semantic mapping with modelling of musical mood values to produce high predictive accuracy of recommendations, in line with a high level of resemblance to biological semantic representations. This would provide personalised recommendations for mood enhancing music that traditional therapies like prescribed playlists could not: the outcome is a form of individualised healthcare, which continually outperforms a more holistic approach (Gürdoğan, Fındık, & Arslan, 2015; Suhonen et al., 2012). The personalisation element is a feature that could also conceivably be assumed to keep depressed listeners with decreased motivation more engaged in robustly mood regulating activity than listening to generic mood enhancing music. Though, in future, this hypothesised technique could be longitudinally compared with traditional methods to reach a substantiated conclusion regarding how long individuals continue with each approach.

To build on the current function of streaming services and their recommender systems, user mood could be extracted using facial or voice assessments of affect from the inbuilt camera or microphone and used to create personalised playlists of track suggestions that follow a trajectory from sadder to more uplifting music. In doing so, the overall result is a form of music content delivery that can provide robust mood enhancement and integrate into people’s lives with minimal impact on daily activity, given the ease of incorporation into commonly used streaming services. Where past solutions would enforce a generic transition from any music that matches a low mood up to more uplifting music, recommender systems could produce dynamic mapping of the semantic and musical features to provide listeners a trajectory through that space to locate personally mood enhancing states. This would build on the initial iso-principle, using the same stepwise approach, but generating solutions that explore the semantic space in ways a human could not and adapting to context in a similar way to biological concept representations. Such a system could potentially allow for improvements in short-term mood levels, if not long-term mental health outcomes, for those who struggle with mood dysregulation. Research on recommender systems could compare the short and long-term outcomes on mood regulation from using either the personalised approach described here or a more traditional method that still uses the iso-principle, thus revealing the validity of applying recommenders to mental health in this manner.

Concluding Remarks

The evidence considered in this dissertation strongly establishes the relevance of music listening habits to emotional wellbeing and justifies several possible methods of integrating a mental health focus into existing music recommendation platforms. The literature confirms a complex and multifaceted relationship between music, listeners and moods, with audio features and lyrical themes inducing powerful emotional states according to culturally specified, or occasionally universal associations. Mapping between elements and emotional valences is shown to be significantly skewed by disorders of mood dysregulation such as depression, and so therapeutic techniques that guide more deliberately mood enhancing music choices have been attempted to mixed results. More consistently effective outcomes have been obtained by following a trajectory that transitions the listener slowly away from mood diminishing sounds in an approach that does not require professional mediation and accounts for decreased motivation in depressed individuals. These principles are well-suited to applications in recommender systems that can continuously recommend music within selected parameters in a way that avoids the impersonal approach of “prescribed playlists”.

Recommender systems function in part by attempting to mimic the semantic structure of the content objects, drawing on empirical evidence that concepts are represented in continuous space and are organised according to similarity. Most mainstream systems fall into the category of distributional models, a type of computational model which only represent concepts according to statistical similarity and therefore produces incomplete representations. Crucial facets of meaning such as compositional and sensorimotor characteristics, as well as context-specified attributes, are ignored. However, for the purpose of personalised music recommendation, LSA appears to be adequate, generating recommendations with impressive predictive accuracy. Precision of such systems has demonstrably increased with the inclusion of contextual factors that drive concept-based decision-making, including location, time and mood. This aligns with the evidence that biological plausibility improves machine learning, as semantic knowledge is context sensitive. Since highly accurate suggestions can be compiled with listener mood factored in, future research should investigate the efficacy of mood-detecting music recommenders on mood regulation outcomes in depressed individuals, following the principles set out above. Studies could focus on areas such as maximising predictive accuracy through proper modelling of mood in music and listener, including elements of meaning that semantic spaces often fail to consider, and utilising longitudinal designs to observe the potential long-term impacts of music-based mood repair.

References

Aalbers, S., Fusar-Poli, L., Freeman, R. E., Spreen, M., Ket, J. C. F., Vink, A. C., Maratos, A., Crawford, M., Chen, X. J. & Gold, C. (2017). Music therapy for depression. *Cochrane Database Of Systematic Reviews*. doi: 10.1002/14651858.CD004517.pub3.

Adomavicius, G., Mobasher, B., Ricci, F. & Tuzhilin, A. (2011). Context-Aware Recommender Systems. *AI Magazine*, *32*(3), 67-80. doi: 10.1609/aimag.v32i3.2364

Alaimo, C. & Kallinikos, J. (2019). Recommender Systems In Beyes, T., Holt, R. & Pias, C. (Eds.) *The Oxford Handbook of Media, Technology and Organization Studies*. Oxford, UK: Oxford University Press.

Andjelkovic, I., Parra, D. & O'Donovan, J. (2016). Moodplay. *Proceedings Of The 2016 Conference On User Modeling Adaptation And Personalization - UMAP '16*. doi: 10.1145/2930238.2930280

APA Dictionary of Psychology. (2020). Retrieved 2 February 2020, from https://dictionary.apa.org/mood

Aucouturier, J. & Pachet, F. (2003). Representing Musical Genre: A State of the Art. *Journal Of New Music Research*, *32*(1), 83-93. doi: 10.1076/jnmr.32.1.83.16801

Azzopardi, J. (2018). Item-Based Vs User-Based Collaborative Recommendation Predictions. In Szymański, J. & Velegrakis, Y. (Eds.) *Semantic Keyword-Based Search on Structured Data Sources. IKC 2017. Lecture Notes in Computer Science*, vol 10546. Cham, Switzerland: Springer.

Balkwill, L. & Thompson, W. (1999). A Cross-Cultural Investigation of the Perception of Emotion in Music: Psychophysical and Cultural Cues. *Music Perception: An Interdisciplinary Journal*, *17*(1), 43-64. doi: 10.2307/40285811

Baltrunas, L., Kaminskas, M., Ludwig, B., Moling, O., Ricci, F., & Aydin, A., Lüke, K. H. & Schwaiger, R. (2011). InCarMusic: Context-Aware Music Recommendations in a Car. *Lecture Notes In Business Information Processing*, 89-100. doi: 10.1007/978-3-642-23014-1_8

Bhatia, S. & Stewart, N. (2018). Naturalistic multiattribute choice. *Cognition*, *179*, 71-88. doi: 10.1016/j.cognition.2018.05.025

Bhatia, S., Zou, W., Richie, R. & Vazire, S. (2019). Predicting High-Level Human Judgment Across Diverse Behavioral Domains. *Collabra: Psychology*, 5(1), 50 doi: 10.1525/collabra.282.

Biswas, R., Riffe, D. & Zillmann, D. (1994). Mood Influence on the Appeal of Bad News. *Journalism Quarterly*, *71*(3), 689-696. doi: 10.1177/107769909407100319

Blattner, M., Zhang, Y. & Maslov, S. (2007). Exploring an opinion network for taste prediction: An empirical study. *Physica A: Statistical Mechanics And Its Applications*, *373*, 753-758. doi: 10.1016/j.physa.2006.04.121

Bonneville-Roussy, A., Rentfrow, P., Xu, M. & Potter, J. (2013). Music through the ages: Trends in musical engagement and preferences from adolescence through middle adulthood. *Journal Of Personality And Social Psychology*, *105*(4), 703-717. doi: 10.1037/a0033770

Brisson, R. & Bianchi, R. (2019). On the relevance of music genre-based analysis in research on musical tastes. *Psychology Of Music*, 1-18. doi: 10.1177/0305735619828810

Burlacu, A. (2018, April 25). Huawei Reportedly Working On AI Digital Assistant That Can Understand Human Emotions, Not Just Commands. Retrieved 3 February 2020, from https://www.techtimes.com/articles/225984/20180425/huawei-reportedly-working-on-ai-digital-assistant-that-can-understand-human-emotions-not-just-commands.htm

Ciocca, S. (2017, October 10). *Spotify's Discover Weekly: How machine learning finds your new music*. Retrieved 2 February 2020, from https://medium.com/s/story/spotifys-discover-weekly-how-machine-learning-finds-your-new-music-19a41ab76efe.

Clark, N. & Tamplin, J. (2016). View of How Music Can Influence the Body: Perspectives From Current Research | Voices: A World Forum for Music Therapy. Retrieved 2 February 2020, from https://voices.no/index.php/voices/article/view/2308/2063

Çukur, T., Nishimoto, S., Huth, A. & Gallant, J. (2013). Attention during natural vision warps semantic representation across the human brain. *Nature Neuroscience*, *16*(6), 763-770. doi: 10.1038/nn.3381

Davenport, C. (2018, April 20). Huawei's new virtual assistant will 'provide emotional interactions'. Retrieved 3 February 2020, from https://www.androidpolice.com/2018/04/20/huaweis-new-virtual-assistant-will-provide-emotional-interactions/

Davidson, J., Liebald, B., Liu, J., Nandy, P., Van Vleet, T., & Gargi, U. (2010). The YouTube video recommendation system. In *Recsys'10: ACM Conf. on Recommender Systems*, 293–296.

Davis, W., Gfeller, K. & Thaut, M. (2008). *An introduction to music therapy*. Silver Spring, Md.: American Music Therapy Association.

Dolan, R. & Sharot, T. (2011). *Neuroscience of Preference and Choice*. Saint Louis, MO: Elsevier Science.

Erdal M., Kächele M. & Schwenker F. (2016). Emotion Recognition in Speech with Deep Learning Architectures. In Schwenker F., Abbas H., El Gayar N., Trentin E. (Eds.) *Artificial Neural Networks in Pattern Recognition. ANNPR 2016. Lecture Notes in Computer Science*, 9896. Cham, Switzerland: Springer

Erkkilä, J., Punkanen, M., Fachner, J., Ala-Ruona, E., Pöntiö, I., Tervaniemi, M., Vanhala, M. & Gold, C. (2011). Individual music therapy for depression: Randomised controlled trial. *British Journal of Psychiatry,* 199(2), 132-139. doi:10.1192/bjp.bp.110.085431

Garrido, S. (2017). *Why are We Attracted to Sad Music?* Cham, Switzerland: Palgrave Macmillan.

Garrido, S. & Schubert, E. (2015a). Moody melodies: do they cheer us up? A study of the effect of sad music on mood. *Psychol. Music* 43, 244–261. doi: 10.1177/0305735613501938

Garrido, S. & Schubert, E. (2015b). Music and People with Tendencies to Depression. *Music Perception: An Interdisciplinary Journal*, *32*(4), 313-321. doi: 10.1525/mp.2015.32.4.313

Garrido, S., Schubert, E. & Bangert, D. (2016). Musical prescriptions for mood improvement: An experimental study. *The Arts in Psychotherapy, 51,* 46–53. doi: 10.1016/j.aip.2016.09.002

Gauthier, I., Behrmann, M. & Tarr, M. (1999). Can Face Recognition Really be Dissociated from Object Recognition?. *Journal Of Cognitive Neuroscience*, *11*(4), 349-370. doi: 10.1162/089892999563472

Grimm, M., Kroschel, K. & Narayanan, S., (2007). Support vector regression for automatic recognition of spontaneous emotions in speech. In *Proc. Internat. Conf. on Acoustics, Speech and Signal Processing*, 4, 1085–1088. doi: 10.1109/icassp.2007.367262

Gürdoğan, E. P., Fındık, U. Y. & Arslan, B. K. (2015). Patients' perception of individualized care and satisfaction with nursing care levels in Turkey. *International Journal of Caring Sciences*, 8, 369–75.

Han, B., Rho, S., Jun, S. & Hwang, E. (2009). Music emotion classification and context-based music recommendation. *Multimedia Tools And Applications*, *47*(3), 433-460. doi: 10.1007/s11042-009-0332-6

Harris, Z. (1970). Distributional structure. In *Papers in structural and transformational Linguistics*, 775–794. Dordrecht, Holland: D. Reidel Publishing Company

HBS Digital Initiative (2018, November 13). Discover Weekly: How Spotify is Changing the Way We Consume Music. Retrieved November 22, 2019, from https://digital.hbs.edu/platform-rctom/submission/discover-weekly-how-spotify-is-changing-the-way-we-consume-music/

Heiderscheit, A. & Madson, A. (2015). Use of the Iso Principle as a Central Method in Mood Management: A Music Psychotherapy Clinical Case Study. *Music Therapy Perspectives*, *33*(1), 45-52. doi: 10.1093/mtp/miu042

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H. & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin*, 31, 1369 –1385. doi: 10.1177/0146167205275613

Hunter P. G., Schellenberg E. G. & Griffith A. T. (2011). Misery loves company: mood-congruent emotional responding to music. *Emotion*, 11, 1068–1072. doi: 10.1037/a0023749

Huth, A., Nishimoto, S., Vu, A. & Gallant, J. (2012). A Continuous Semantic Space Describes the Representation of Thousands of Object and Action Categories across the Human Brain. *Neuron*, *76*(6), 1210-1224. doi: 10.1016/j.neuron.2012.10.014

Jentschke, S. (2007). *Neural correlates of processing syntax in music and language – influences of development, musical training, and language impairment.* (Unpublished

Doctoral Thesis) Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig

Johnsrude, I., Penhune, V. & Zatorre, R. (2000). Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain*, *123*(1), 155-163. doi: 10.1093/brain/123.1.155

Jones, M. N., Willits, J. A. & Dennis, S. (2015). Models of semantic memory. In Busemeyer, J. R., Wang, Z., Townsend, J. T. & Eidels, A. (Eds.) *Oxford Handbook of Mathematical and Computational Psychology* (pp. 232–54). Oxford, UK: Oxford University Press.

Kawakami, A. (2013, September 20). Opinion | Why We Like Sad Music. Retrieved 2 February 2020, from https://www.nytimes.com/2013/09/22/opinion/sunday/why-we-like-sad-music.html

Keil, F. C. (1979). *Semantic and conceptual development: An ontological perspective*. Cambridge, MA: Harvard University Press.

Kessler, R. & Bromet, E. (2013). The Epidemiology of Depression Across Cultures. *Annual Review Of Public Health*, *34*(1), 119-138. doi: 10.1146/annurev-publhealth-031912-114409

Kim, J., Demetriou, A., Manolios, S. & Liem, C. (2019). Beyond Explicit Reports. *Proceedings Of The 27Th ACM Conference On User Modeling, Adaptation And Personalization - UMAP '19*. doi: 10.1145/3320435.3320462

Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Current Opinion In Neurobiology*, *15*(2), 207-212. doi: 10.1016/j.conb.2005.03.005

Li, X., Hong, X., Moilanen, A., Huang, X., Pfister, T., Zhao, G. & Pietikainen, M. (2017). Towards Reading Hidden Emotions: A Comparative Study of Spontaneous Micro-

Expression Spotting and Recognition Methods. *IEEE Transactions On Affective Computing*, *9*(4), 563-577. doi: 10.1109/taffc.2017.2667642

Macario, J. (1991). Young children's use of color in classification: Foods and canonically colored objects. *Cognitive Development*, *6*(1), 17-46. doi: 10.1016/0885-2014(91)90004-w

Manuel, P. (2005, October 14). Does sad music make one sad? An ethnographic perspective. *Contemporary Aesthetics,* 3. Retrieved 2 February 2020 http://www.contempaesthetics.org/newvolume/pages/article.php?articleID=286

Marinkovic, K., Dhond, R., Dale, A., Glessner, M., Carr, V. & Halgren, E. (2003). Spatiotemporal Dynamics of Modality-Specific and Supramodal Word Processing. *Neuron*, *38*(3), 487-497. doi: 10.1016/s0896-6273(03)00197-1

Matt, G., Vázquez, C., & Campbell, W. (1992). Mood-congruent recall of affectively toned stimuli: A meta-analytic review. *Clinical Psychology Review*, *12*(2), 227-255. doi: 10.1016/0272-7358(92)90116-p

McFerran, K. S., Hense, C., Koike, A. & Rickwood, D. (2018). Intentional music use to reduce psychological distress in adolescents accessing primary mental health care. *Clinical Child Psychology and Psychiatry*, *23*(4), 567–581. doi: 10.1177/1359104518767231

Meadowcroft, J. & Zillmann, D. (1987). Women's Comedy Preferences During the Menstrual Cycle. *Communication Research*, *14*(2), 204-218. doi: 10.1177/009365087014002004

Meehan, K., Lunney, T., Curran, K. & McCaughey, A. (2013). Context-aware intelligent recommendation system for tourism. *2013 IEEE International Conference On Pervasive Computing And Communications Workshops (PERCOM Workshops)*. doi: 10.1109/percomw.2013.6529508

Meng, X., Wang, S., Liu, H. & Zhang, Y. (2018). Exploiting emotion on reviews for recommender systems. In *AAAI*.

Mervis, C. B. (1987). Child-basic object categories and early lexical development. In Neisser, U. (Ed.), Concepts and conceptual development: Ecological and intellectual factors in categorization (pp. 201- 233). Cambridge, England: Cambridge University Press.

Mollo, G., Cornelissen, P., Millman, R., Ellis, A. & Jefferies, E. (2017). Oscillatory Dynamics Supporting Semantic Cognition: MEG Evidence for the Contribution of the Anterior Temporal Lobe Hub and Modality-Specific Spokes. *PLOS ONE*, *12*(1). doi: 10.1371/journal.pone.0169269

North, A., Hargreaves, D. & Hargreaves, J. (2004). Uses of Music in Everyday Life. *Music Perception: An Interdisciplinary Journal*, *22*(1), 41-77. doi: 10.1525/mp.2004.22.1.41

Pagano, R., Cremonesi, P., Larson, M., Hidasi, B., Tikk, D., Karatzoglou, A. & Quadrana, M. (2016). The Contextual Turn. *Proceedings Of The 10Th ACM Conference On Recommender Systems - Recsys '16*, 249-252. doi: 10.1145/2959100.2959136

Patterson, K., Nestor, P. & Rogers, T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, *8*(12), 976-987. doi: 10.1038/nrn2277

Pfister, T., Li, X., Zhao, G. & Pietikainen, M. (2011). Recognising spontaneous facial micro-expressions. *2011 International Conference On Computer Vision*. 1449-1456. doi: 10.1109/iccv.2011.6126401

Platt, B., Murphy, S. & Lau, J. (2015). The association between negative attention biases and symptoms of depression in a community sample of adolescents. *PeerJ*, *3*, e1372. doi: 10.7717/peerj.1372

Polanco-Roman, L., Jurska, J., Quiñones, V. & Miranda, R. (2015). Brooding, Reflection, and Distraction: Relation to Non-Suicidal Self-Injury versus Suicide Attempts.

Archives of Suicide Research: Official Journal of the International Academy for Suicide Research, 19(3), 350-365. doi: 10.1080/13811118.2014.981623.

Prey, R. (2018). Nothing personal: algorithmic individuation on music streaming platforms. *Media, Culture & Society*, *40*(7), 1086–1100. doi: 10.1177/0163443717745147

Raes, F., Hermans, D., Williams, J., Beyers, W., Eelen, P. & Brunfaut, E. (2006). Reduced autobiographical memory specificity and rumination in predicting the course of depression. *Journal Of Abnormal Psychology*, *115*(4), 699-704. doi: 10.1037/0021-843x.115.4.699

Ralph, M., Jefferies, E., Patterson, K. & Rogers, T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, *18*(1), 42-55. doi: 10.1038/nrn.2016.150

Rentfrow, P. & Gosling, S. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal Of Personality And Social Psychology*, *84*(6), 1236-1256. doi: 10.1037/0022-3514.84.6.1236

Reskinoff, P. (2014, June 25). Google Executive: "You Cannot Devalue Music. It's Impossible.." Retrieved 3 February 2020, from https://www.digitalmusicnews.com/2014/06/25/devaluationmyth/

Rips, L., Shoben, E. & Smith, E. (1973). Semantic distance and the verification of semantic relations. *Journal Of Verbal Learning And Verbal Behavior*, *12*(1), 1-20. doi: 10.1016/s0022-5371(73)80056-8

Rogers, T. & McClelland, J. (2004). *Semantic cognition*. Cambridge, MA: MIT Press.

Rosch, E. (1973). Natural categories. *Cognitive Psychology*, *4*(3), 328-350. doi: 10.1016/0010-0285(73)90017-0

Ruiz-Montiel, M., & Aldana-Montes, J. (2009). Semantically Enhanced Recommender Systems. *Lecture Notes In Computer Science*, 604-609. doi: 10.1007/978-3-642-05290-3_74

Rumelhart, D. E. (1990). Brain style computation: Learning and generalization. In Zornetzer, S. F., Davis, J. L. & Lau, C. (Eds.), *An introduction to neural and electronic networks*, 405-420. San Diego, CA: Academic Press.

Rutter, M. (2012). Resilience as a dynamic concept. *Dev. Psychopathol.* 24, 335–344. doi: 10.1017/S0954579412000028

Saarikallio, S. & Erkkilä, J. (2007). The role of music in adolescents' mood regulation. *Psychology Of Music*, *35*(1), 88-109. doi: 10.1177/0305735607068889

Sakka, L., & Juslin, P. (2018). Emotion regulation with music in depressed and non-depressed individuals. *Music & Science*, *1*, 1-12. doi: 10.1177/2059204318755023

Schäfer, T. & Mehlhorn, C. (2017). Can personality traits predict musical style preferences? A meta-analysis. *Personality And Individual Differences*, *116*, 265-273. doi: 10.1016/j.paid.2017.04.061

Schramm, H., & Cohen, E. (2017). Emotion Regulation and Coping via Media Use. *The International Encyclopedia Of Media Effects*, 1-9. doi: 10.1002/9781118783764.wbieme0162

Scott, S., Blank, C., Rosen, S. & Wise, R. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, *123*(12), 2400-2406. doi: 10.1093/brain/123.12.2400

Seo, Y. & Huh, J. (2019). Automatic Emotion-Based Music Classification for Supporting Intelligent IoT Applications. *Electronics*, *8*(2), 164. doi: 10.3390/electronics8020164

Sha, L., Haxby, J., Abdi, H., Guntupalli, J., Oosterhof, N., Halchenko, Y. & Connolly, A. (2015). The Animacy Continuum in the Human Ventral Vision Pathway. *Journal Of Cognitive Neuroscience*, *27*(4), 665-678. doi: 10.1162/jocn_a_00733

Shardanand, U. & Maes, P. (1995). Social information filtering: Algorithms for automating 'word of mouth'. *Proceedings Of The SIGCHI Conference On Human Factors In Computing Systems - CHI '95*. doi: 10.1145/223904.223931

Solomon, S. H., Medaglia, J. D. & Thompson-Schill, S. L. (2019). Implementing a concept network model. *Behav Res* 51**,** 1717–1736. doi: 10.3758/s13428-019-01217-1

Suhonen, R., Papastavrou, E., Efstathiou, G., Tsangari, H., Jarosova, D., & Leino-Kilpi, H., Patiraki, E., Karlou, C., Balogh, Z. & Merkouris, A. (2012). Patient satisfaction as an outcome of individualised nursing care. *Scandinavian Journal Of Caring Sciences*, *26*(2), 372-380. doi: 10.1111/j.1471-6712.2011.00943.x

Sumare, S. P. & Bhalke, M. D. (2015). Automatic Mood Detection of Music Audio Signals: An Overview. *IOSR Journal of Electronics and Communication Engineering*, 83- 87.

Swaminathan, S. & Schellenberg, E. (2015). Current Emotion Research in Music Psychology. *Emotion Review*, *7*(2), 189-197. doi: 10.1177/1754073914558282

Taruffi, L. & Koelsch, S. (2014). The Paradox of Music-Evoked Sadness: An Online Survey. *Plos ONE*, *9*(10), e110490. doi: 10.1371/journal.pone.0110490

ter Bogt, T., Canale, N., Lenzi, M., Vieno, A. & van den Eijnden, R. (2019). Sad music depresses sad adolescents: A listener's profile. *Psychology Of Music*. doi: 10.1177/0305735619849622

Thompson-Schill, S. L., Aguirre, G. K., Desposito, M. & Farah, M. J., (1999). A neural basis forcategory and modality specificity of semantic knowledge. *Neuropsychologia*, *37*(6), 671-676. doi: 10.1016/s0028-3932(98)00126-2

Truschel, J. (2019, November 25). Depression Definition and DSM-5 Diagnostic Criteria. Retrieved 2 February 2020, from https://www.psycom.net/depression-definition-dsm-5-diagnostic-criteria/

Vandenberghe, R., Price, C., Wise, R., Josephs, O. & Frackowiak, R. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature*, *383*(6597), 254-256. doi: 10.1038/383254a0

Vuoskoski, J. K. & Eerola, T. (2012). Can sad music really make you sad? Indirect measures of affective states induced by music and autobiographical memories. *Psychology of Aesthetics, Creativity, and the Arts, 6*(3), 204–213. doi: 10.1037/a0026937

Wang, C., Wang, Y. & Chou, S. (2018). A context and emotion aware system for personalized music recommendation. *Journal of Internet Technology*, 19, 765-779. doi: 10.3966/160792642018051903013.

Warrington, E. (1975). The Selective Impairment of Semantic Memory. *Quarterly Journal Of Experimental Psychology*, *27*(4), 635-657. doi: 10.1080/14640747508400525

Watson, A. (2019, November 20). Spotify users by age in the U.S. 2018 | Statista. Retrieved 2 February 2020, from https://www.statista.com/statistics/475821/spotify-users-age-usa/

Yamins, D., Hong, H., Cadieu, C., Solomon, E., Seibert, D., & DiCarlo, J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proceedings Of The National Academy Of Sciences*, 111(23), 8619-8624. doi: 10.1073/pnas.1403112111

Zatorre, R. & Salimpoor, V. (2013). From perception to pleasure: Music and its neural substrates. *Proceedings Of The National Academy Of Sciences*, *110*(2), 10430-10437. doi: 10.1073/pnas.1301228110