Music & Science
XX(X):1–18

DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

Polarization through Colombian not-so-popular music and algorithms: appraisal guided musically induced emotions

Juan Sebastián Gómez-Cañón[1], Thomas Magnus Lennie[2], Tuomas Eerola[2], Pablo Aragón[3], Estefanía Cano[4], Perfecto Herrera[1], Emilia Gómez[1,5]

Abstract

This work presents a study of how Music Emotion Recognition (MER) systems could be biased with respect to annotations of musically-induced emotions in a political context. Specifically, we analyze traditional Colombian music containing politically charged lyrics of two types: (1) vallenatos and social songs from the "left-wing" guerrilla Fuerzas Armadas Revolucionarias de Colombia (FARC) and (2) corridos from sympathizers of the "right-wing" paramilitaries Autodefensas Unidas de Colombia (AUC). We train personalized machine learning models to predict induced emotions for participants with diverse political views – we aim at identifying the songs that may induce negative emotions for a particular user, such as anger and fear. By acknowledging the context of the user (their political leaning), we show that user's emotion judgments could be interpreted as problematizing data – subjective emotional judgments could in turn be used to influence the user in a human-centered machine learning environment. In short, highly desired "emotion regulation" applications could potentially deviate to "emotion manipulation" – the recent discredit of emotion recognition technologies might transcend ethical issues of diversity and inclusion.

Keywords

personalization, polarization, music emotion recognition

Introduction

Emotion is acknowledged as one of the core reasons people engage with music (Juslin et al. 2015). Yet, the mechanisms that underpin the emotional effects of music are still a matter of debate (Céspedes-Guevara and Eerola 2018; Warrenburg 2020). We are likely still some time from a comprehensive model of musical emotions. Nonetheless, progress has been made in recent attempts by Lennie and Eerola (2022) to situate emotional responses to music within contemporary developments in the affective and cognitive sciences – the CODA Model: Constructivistly-Organised Dimensionsal-Appraisal. Lennie and Eerola (2022) hypothesize that affective responses to music, like many utilitarian affective responses (Frijda 2007; Moors 2017), are driven by goal-directed mechanisms. Specifically, the CODA model hypothesizes that an individual's goals and the relevance of a stimulus to those goals in a specific situation will influence the development of an emotional episode induced by music. Moreover, this relationship between goal-directed mechanisms and music is predicted to be bi-directional. That is, a musical stimuli seen as relevant to one's goals may amplify the degree of relevance that an individual places on a particular goal. Similarly, a stimuli that is in conflict with an individual's goals may amplify negative reactions through an inability to achieve or align with one's goals. For example, a song that discredits Apple devices could be used with different purposes as it could yield different reactions (and possibly different induced emotions) – depending on the congruence of the individual's goals and the relevance of the song. The value of musical goals has been acknowledged in previous literature (Sloboda and Juslin 2010) and has been included in several theoretical models (Scherer and Coutinho 2013; Céspedes-Guevara 2021). Some models explicitly hypothesize the link between goal-directed mechanisms and core-affect (Thompson et al. 2011; Lennie and Eerola 2022).* Yet, little empirical work has followed, possibly due to a historical narrative that aesthetic emotions, including musical emotions, have little influence on life goals (Kant 1790). This is an idea Huron (2016) describes as irreconcilable with current biological understanding (p. 242). This goal-directed and context dependent understanding of musical emotions allows for substantial individual variation in emotional responses to music. Subsequently, it provides a valuable theoretical starting point for explaining individual differences in Music Emotion Recognition (MER).

[1]Music Technology Group, Universitat Pompeu Fabra, Spain
[2]Department of Music, Durham University, UK
[3]Universitat Pompeu Fabra, Spain
[4]Songquito, Germany
[5]European Commission, Joint Research Centre, Spain

Corresponding author:
J.S. Gómez-Cañón Music Technology Group, Universitat Pompeu Fabra, Barcelona, 08002, Spain
Email: juansebastian.gomez@upf.edu

*Core affects are arousal and valence. Arousal refers to energy or activation and valence relates to pleasantness or positiveness of an emotion.

From the computational perspective, MER attempts to predict the emotion perceived by or induced in a particular listener (Yang and Chen 2011). To design predictive models it is necessary to obtain a "ground truth" – a term that refers to the "real" or "true" information that the machine learning algorithm attempts to predict (Schuller 2013). A main limitation to MER is to attempt to create this "ground truth" due to the subjectivity of the task. However, a growing effort has been made to produce enriched datasets of emotion judgments with more listening data to better represent the properties and context of the listener (Barthet et al. 2013; Schedl et al. 2013; Gómez-Cañón et al. 2021a): demographics, cultural and individual differences, preference, familiarity, functional uses of music, physiological signals, and language. In this context, personalized models which incorporate this information could more easily predict the particular emotion judgments from a particular listener (Yang et al. 2007; Su and Fung 2012; Chen et al. 2014; Gómez-Cañón et al. 2021b). However, only recently research has been made to study the impact of "toying with emotions combined with personalization" in a decision-making process due to increasing concerns regarding dark patterns and manipulative personalization in digital commercial practices (European Commission et al. 2022) – MER systems could potentially be harmful when inducing particular emotions to a listener. One way in which this could occur is through the manipulation of goal-directed mechanisms, where several aspects of musical engagement may be negatively affected by the misuse of such systems. For example, social functions of music such as group identity or political stance can be publicly misrepresented in playlist recommendations. Moreover, personalization strategies, which should target societal well-being as a requirement to produce trustworthy artificial intelligence[†], are frequently imposing unbalanced and unfair digital asymmetries onto more vulnerable societies in forms of "colonial value and power paradigms" that researchers should strive to acknowledge (Mohamed et al. 2020; Adams 2021; Birhane and Guest 2021). From the field of music technology, Huang et al. (2021) have pointed out how ethical concepts such as "human rights", "well-being" and "potential misuse" – that are typically used in Western societies – need to be carefully examined in other cultural contexts.

Aims

This study is an extension from the work presented by Gómez-Cañón et al. (2021) and Lennie and Eerola (2022) – the aim of this study is to attempt to understand if an MER algorithm can effectively be biased to classify music which can induce "negative" emotions to a listener through the manipulation of goal-directed mechanism. In essence, we wish to empirically show the validity of goal-directed mechanisms in musically induced emotions and how this directly relates to real-world applications of music listening. We use the word "negative" carefully, since the subjectivity of the task complicates defining if an emotion is negative or not. However, a personalized MER algorithm can produce music recommendations grounded on the listeners' judgments (already standard in streaming platforms), which in turn could prove beneficial or harmful to this listener when inducing particular emotions.

This study diverges from previous research which aims at using music for beneficial purposes: to enhance memory, relieve boredom, improve concentration, promote prosociality, or aid learning (Hu et al. 2021; Agres et al. 2021; Cespedes-Guevara and Dibben 2021). The reason is that, despite the common consensus regarding beneficial uses of music, it has also aided to create generalized misconceptions – for example, the Mozart effect (Mehr et al. 2013), binaural beats (Orozco Perez et al. 2020), or 432 Hz tuning (Rosenberg 2021). Only recently has research started to theorize and analyze music-induced harm (Ziv 2016; Silverman et al. 2020) – a topic that should be more widely studied by academia. Moreover, social networks, streaming platforms, and personalized ad companies are already making use of emotional responses to maximize users' engagement (O'Neil 2016; Noble 2018; Zuboff 2019; Véliz 2020). The episode with Cambridge Analytica demonstrated that technology can even persuade users, polarize opinions, and affect decision-making processes by promoting/manipulating emotional stimuli. This form of persuasion, where individuals appear immune to any evidence contrary to their own views, has colloquially been referred to as the "filter bubble" Pariser (2011) or "echo chamber" effect (Sunstein 2002; Garrett 2009), but most notably refers to the greater presentation of information amenable to an individual's existing view points. The homophily principle argues that "similarity breeds connection" (Mcpherson et al. 2001) – personal networks tend to be heterogeneous regarding demographic, behavioral, and individual properties. Thus, studies have evaluated how the interaction of homogeneous social networks in social media, along with disinformation strategies can conduce to political polarization. We refer the reader to Tucker et al. (2018) for a thorough literature review on this topic. More recently, the impact of algorithms in polarization have been questioned in order to produce systems that "depolarize by design" (Stray 2021; Fabbri et al. 2022). However, most of these studies have been carried out using very direct stimuli and interactions between users of social media (e.g., online political conversations on Twitter).

Thus, we present an experiment which addresses polarization in a two-fold approach: (1) by using music (and specifically lyrics) as a emotional stimuli to users, and (2) by using personalization algorithms that attempt to predict the emotion induced by the music to the listener. The Colombian presidential elections offered a unique opportunity to access a time where listeners would show strong political opinions (goal-relevance) and strong emotional responses. The contextualization of the Colombian political landscape escapes the scope of this paper, thus we refer the reader to Chomsky (2004); Zamosc (1986); Berquist (1978); Stokes (2005); Arocha R. et al. (1988); Fals Borda et al. (2001); Mahoney (2020) for deeper analysis regarding the history of violence in Colombia. In a broad context, the "biblical

[†] https://op.europa.eu/s/pInE

holocaust" of Colombian violence – portrayed by the writer Gabriel García Márquez – has resulted in more than 420,000 violent deaths over the last 70 years, more than 11 million Colombians leaving the country or internally displaced, and one of the most unequal distribution of income in the continent (Mahoney 2020). Diverse sources of inequality (i.e., agrarian capitalism, socio-economic exclusion, decolonization processes, the war on drugs, illegal economies, and exploitation of natural resources) are the cause of the formation of illegal armies fueled by political ideologies (Grajales 2021): "left-wing" Fuerzas Armadas Revolucionarias de Colombia (FARC) and "right-wing" Autodefensas Unidas de Colombia (AUC), amongst several other illegal groups. As an oversimplification of Colombia's historical process (and reflecting the generalized trend in the world), polarization has arisen over whether and how to pursue peace in the country, producing negative relationships between political discourses and everyday life (Feldmann 2019).
Music has been used as political propaganda reflecting this polarization of strong left/right political stances in Colombia. Music genres have been associated with particular political ideologies. The "left-wing" guerrilla group FARC is associated with vallenatos[‡] and social songs that support their political ideology. Similarly, the corridos prohibidos – ballads often associated with the 20th century Mexican revolution and later with the narcotics trade (Villalobos and Ramírez-Pimienta 2004; Barbosa Caro and Suavita 2019) – have been used by several sympathizers of the "right-wing" paramilitaries AUC. While neither of the types of music is popular by Colombian chart standards, their political allegiance is clear within the lyrical content and the political stances they support are well understood in the population – we have previously referred to this music as Colombian not-so-popular music (Gómez-Cañón et al. 2021).

Through the interdisciplinary link between affective cognition and MER, we formulate and study the following research questions:

- R1 Do an individual's political values (goals) influence core-affect in emotional episodes induced by music?
- R2 Can a MER algorithm for induced emotions be biased towards a particular opinion with respect to music with polarizing lyrics?

Hypotheses

- H1 Participants with different political viewpoints (goal-relevance) will show different induced arousal annotations.
- H2 Politically sensitive music stimuli that agree/disagree with participants political stance (goal-congruence) will show different induced valence annotations.
- H3 Personalized algorithms will be effectively biased towards specific categories depending on the political stance of listeners – namely, personalized algorithms reflect the listeners' political views.

Contributions

- We empirically test the validity of goal-directed mechanisms in musically induced emotions.
- We evaluate how personalization of MER algorithms can disclose sensitive personal data.

Related work

A goal-directed approach

A goal-directed mechanism can be described as cognitive processes that relate to an organisms needs, wishes, desires, values or beliefs, grouped together broadly as goals. Goals can be described at different functional levels. For example, lower-level goals included basic survival functions like food or sex, while higher-level goals include higher cognitive functioning such as personal values or social identity, sometimes described as "norm-compatibility". A goal-directed mechanism is more typically described as a competition between multiple competing goals (Moors 2017).

Goal-directed processes have become a central part of the wider research in to the affective and cognitive sciences (Sloman 1996; Balleine and Dickinson 1998; Frijda 2007; Eder and Hommel 2013; Schiller 2022). The most popular music emotion model (Juslin and Västfjäll 2008), however, makes no mention of goals and has been heavily criticised for this (Moors and Kuppens 2008; Scherer and Zentner 2008; Scherer and Coutinho 2013; Céspedes-Guevara 2021; Lennie and Eerola 2022). Later variations of this model (Juslin 2013, 2019) acknowledge that goals are present but distinguishes them as separate from other underlying mechanisms and describes them as occurring "rarely" (Juslin 2013, p. 239). In contrast, Lennie and Eerola (2022) have proposed a new model that seeks to re-center goal-directed processes into the mechanisms that underpin emotional responses to music. Uniquely, they draw together to competing emotion models: Constructionist (Russell 1980, 2003; Barrett 2006, 2017) and Dimensional-Appraisal (Ellsworth and Scherer 2003; Scherer 2009a,b; Moors et al. 2013) through goal-directed processes, reflecting recent moves in the affective sciences (Schiller 2022).

Some goal-directed theories have made explicit predictions about how different goals may be prioritized over others in different context (Moors et al. 2017) and drive further attention and cognitive resources towards this goal. Put simply, certain contexts may lead to the prioritization of certain goals over others. For instance, during political elections people may give greater priority to evaluating their political position or may invest greater cognitive resources towards defending or strengthening their existing political position; be it intentionally or unintentionally – the "echo chamber" phenomenon (Garrett 2009). Explicit predictions about how stimuli interact with an organism's goals leading to changes in other components of emotion have been made (Moors et al. 2013). Evidence for this has come from studies showing that multiple appraisal dimensions, including goal-relevance and goal-congruence (how well a stimulus supports or inhibits a goal), can bi-directional interact and modify core-affect (Kuppens et al. 2012).

[‡] A Colombian folk tradition that translates as "born in the valley".

This interaction is often closely related to how congruent or incongruent a stimulus is for an organism achieving its goals. More recently, studies testing the causal nature of these mechanisms over time have shown that goal-relevant stimuli lead to greater changes in core-affect compared to irrelevant ones, most prominently in experienced pleasantness (Asutay and Västfjäll 2021).

Goal-directed processes have long been associated with personal and social values (Frijda et al. 1986; Frijda 2007; Smith and Lazarus 1991; Scherer 1982, 2009a; Ellsworth and Scherer 2003). The evidence further suggest that stimuli that are congruent or incongruent with an individual's personal identity or social values can also influence the development of an emotional episode (Scherer and Moors 2019, i.e., influencing other components of emotion).

Social & personal identities

Music preferences are a commonly cited feature in the formation of social identities (Tarrant et al. 2002; Shepherd and Sigg 2015) and personal values (Steele and Brown 1995; Saarikallio 2019). In this sense, music is viewed as a resource in the development of these components rather than something that directly influences them. Evidence for music as a mediating influence in the development of these components has emerged. Tarrant (2002) has shown that preferences for particular music may lead to attributing similar personality traits to individuals that are associated with the social group as a whole. Under the umbrella of social and personal identities, research has provided firm grounding for the correlation of music preferences with personality traits, value systems, cognitive abilities, perceptions of gender (Zweigenhaft 2008; Dunn et al. 2012; Delsing et al. 2008; Ter Bogt et al. 2010) and critically political ideologies Rentfrow and Gosling (2003); Rentfrow et al. (2011); Carney et al. (2008), see Rentfrow (2012) for a general review. Studying a sample of American participants, Carney et al. (2008, Study 3) found that conservative viewpoints correlated with more conventional music choices and owning fewer variety of CD's. Liberals alternatively, showed greater preference for world, folk, classical, contemporary rock, and golden oldies. Unsurprisingly, this result similarly correlates with the personality dimension "openness to experience". A somewhat cautious interpretation of Carney et al. (2008) is warranted here due to the large number of statistical tests conducted. This has possibly led to several chance findings. In a similar vein, Rentfrow et al. (2011) found that "aesthetic and complex" entertainment types, and in Rentfrow and Gosling (2003) "reflective and complex" and "energetic and rhythmic" music, correlated with a liberal self-identity. This finding was negatively correlated with social dominance orientation (Rentfrow and Gosling 2003). Alternatively, and in agreement with Carney et al. (2008), Rentfrow and Gosling (2003) found self-identified right-leaning participants preferred "upbeat and conventional" music. Of interest to our hypotheses, Rentfrow and Gosling (2003) note that future directions should more openly explore the like between music preferences with values and goals (p. 1251).

Other findings (Heisbourg and Feitosa 2021) have linked music to marginal influences on the perception of some personality traits in political representatives in a positive direction. However, it remains to be shown if music can negatively influence perceptions of certain personality traits. The Heisbourg and Feitosa study also made no assessment of participants' preference for the presented music, its lyrical content, or their general music preferences. Furthermore, the study's focus on influencing a participant's perception of another persons personality is not directly related to the participants' personal political stance or their emotional response to the stimuli. This makes it difficult to relate the already marginal effects directly to the music literature. For example, does music that is incongruent with an individual's personal values or social identity, such as political ideologies, induce negative emotions (i.e., fear, anger, disbelief, sadness)?

Given the wealth of literature associating music preferences with personal identity and social values, including political orientation (Rentfrow et al. 2011; Carney et al. 2008), it seems prudent to hypothesize that music can be used as a stimulus that can be appraised as either congruent or incongruent with an individual's goals in a political context. Yet, no musical study of this kind has been attempted to the authors' knowledge.

Music emotion recognition

Music emotion recognition (MER) is a computational task that evaluates emotionally relevant features from music and correlates them with certain emotions, be it *perceived* or *induced* in a listener. Interestingly, while we may *perceive* that certain music expresses a "happy" emotion, since it is upbeat and in major tonality for a Western context, it does not necessarily makes us *feel* "happy" (Gabrielsson 2006). This is an important distinction to understand the complexity of producing a "ground truth" for MER systems since low quality annotations have a direct impact on the overall performance of the algorithms (Hsueh et al. 2009). MER uses features that are typically low-level acoustic representations of sound (e.g., tempo or pitch), high-level semantic descriptions (e.g., genre), information extracted from lyrics, and data about the listeners' properties or context (e.g., collecting physiological data from listeners Hu et al. (2018)). Machine learning algorithms are then used to correlate these features with an emotion "ground truth" (Laurier and Herrera 2009; Laurier 2011; Yang and Chen 2011) – emotion judgments are typically collected through subjective listening tests in which listeners annotate excerpts of music.

Panda et al. (2020) recently reviewed several emotionally-relevant acoustic features. For example, melody relates to fundamental frequency f_0 or pitch salience; rhythm relates to note onsets or note durations; dynamics relate to sound level or note intensity; timbre relates to spectral centroid or mel-frequency cepstral coefficients. Towards a more multi-modal approach, da Silva Mahleiro (2016) combined acoustic features and data extracted from lyrics to predict the emotions in music – his studies reveal up to 9 percentage point improvement of F-scores by using a multi-modal classification approach. Textual features may be extracted from lyrics through natural language processing methods and can be used to analyze sentiment from text (Manning and Schütze 1999; Jurafsky and Martin 2009) – language processing has been at the heart of machine learning by extracting meaningful information from text. Recently, Agrawal et al. (2021) used novel deep neural networks to predict emotions in lyrics. The use of this information can enhance the performance of a multi-modal MER system.

Given the subjectivity of the annotation task in producing a “ground truth”, Yang et al. (2007) proposed that the response variability of each individual listener could be better modeled by introducing personalized models – a personalized model is trained exclusively from the annotations of an individual listener. Sarasúa et al. (2012) and Su and Fung (2012) leveraged active learning to improve MER performance and produce personalized models, respectively – active learning is based on selecting specific training instances such that algorithms perform better with less training. Gómez-Cañón et al. (2021b) proposed consensus entropy to improve active learning by leveraging the collective judgments from music – the main assumption is that the data instances, which a collective of annotators disagrees upon, reflect individual boundaries that could be informative for personalization. The TROMPA-MER dataset was created with a human-centric approach for data collection and annotation (Gómez-Cañón et al. 2022): listeners annotated each music excerpt with single free-text emotion words (in native language), distinct forced-choice emotion categories, preference, and familiarity. In short, MER has started to incorporate more personal and contextual data as a way to improve the performance of the system – this type of data becomes central to the future of MER and potential applications that will follow in the future.

From the perspective of ethics in artificial intelligence, personal data (personal interests, preferences, psychological profile, mood) and the growth of machine learning systems (involving personalization practices) have been typically used for persuasion purposes – steering, coercing or manipulating users into making decision that may or may not be in their best interest (European Commission et al. 2022). There is a need to evaluate the role of data as a resource typically exploited for economic expansion (Mohamed et al. 2020) – artificial intelligence may conceal asymmetrical power relations that are difficult to assess by developers. The notion of algorithmic oppression is generally linked to decision systems with direct impact on users: linking criminal datasets to discriminatory police practices (O'Neil 2016), using facial recognition systems that fail to recognize Black faces (Buolamwini and Gebru 2018), and threatening privacy and enforcing racist pseudo-sciences by using facial emotion recognition systems (Stark 2016; Crawford 2021). Conversely, the apparently innocuous task of MER has only recently shown the advantages of gathering data such as listening habits or physiological signals (Hu et al. 2018; Gómez-Cañón et al. 2021a) – the potential ethical issues of MER will appear as sensitive emotional data is processed by personalization systems that increasingly improve with the users' engagement. Thus, and to the knowledge of the authors, MER systems have not been evaluated to this extent in the past.

Methodology

Experimental setup

The experiment was held in an online platform developed using React and a Flask backend. It was developed in English (for design purposes) and translated to Spanish (for the participants)[§]. The experiment was first piloted with 5 Colombian nationals. Participants were collected through social media outlets (Twitter and Instagram) targeting Colombian nationals over the course of 2 months (June-July) in the run up to the 2022 presidential elections and after the first round of elections. We also invited researchers from the following Colombian universities: Universidad Central, Universidad de Antioquia, Universidad de los Andes, and Universidad Nacional. Participants were incentivized through a prize draw for one of four €100 vouchers.

We summarize our experimental procedure as follows:

1. Participants agreed to the consent form validated by the ethics committee from the Universitat Pompeu Fabra (in conformity with GDPR and Colombian data protection laws) – stating that they are Colombians and are between 18 and 65 years old.
2. Participants provided general demographic variables: age, gender, native language, and musical self-identification.
3. A short attention check followed: three sounds in randomly assigned order were played and each participant should select the one with lowest sound level. If the participant was not able to complete this step, they were informed that their device did not have the audio fidelity to continue with the experiment.
4. Participants received an explanation on how to complete the annotations: a comparison between perceived and induced emotions in music and a description of the annotation interface (see section on annotation gathering).
5. Participants carried on to annotate the music: (a) each participant was randomly assigned a different personaization strategy: a model trained with acoustic features (*ACO*), a model trained with features from the lyrics (*LYR*), a multi-modal model that takes into account both lyrics and acoustic features (*MIX*), and a pseudo-random baseline that presented music to perform an annotation consistency check (*RAND*); (b) initially, all participants annotated the same 6 tracks for the first iteration (2 FARC-songs; 2 AUC-songs; 2 songs without lyrics which were randomly selected) to train personalized models (see section on personalization); (c) based on the output of the first annotations, the personalized model was retrained and queried a new batch of 6 tracks to be annotated by the participant; and (d) the personalized model was then refined using the remaining tracks presented in

[§] https://trompa-mtg.upf.edu/colombian-not-popular/

4 iterations of 6 tracks each (i.e., each participant annotated 30 tracks in total).

6. Finally, participants completed three questionnaires on the political opinion – Right-Wing Authoritarianism scale (RWA), the Social Dominance Orientation scale (SDO), and a Colombian specific political questionnaire made the purposes of this study.

The experiment took an average of 30-40 minutes to complete.

Scales

Musical self-identification was measured by the Ollen Musical Sophistication Index (OMSI). This single item scale provides estimates of the psychometric measures used to determine membership of the category "musician" (Ollen 2006). The OMSI musician rank item is concerned with the individual's self-assessed level of musical identity as opposed to an item relating to musical expertise (Zhang and Schubert 2019).

The *right-wing authoritarianism* (RWA) scale, originally developed by Altemeyer (1998), measures social attributes, such as the degree to which people defer to established authorities, show aggression toward out-groups when authorities sanction that aggression, and support traditional values endorsed by authorities (Saunders and Ngo 2017), racism and sexism (Zakrisson 2005). We used the short 15-item version of the RWA (Zakrisson 2005) which uses less extreme and more modern language, and makes less reference to specific groups (e.g., women).

Social dominance orientation (SDO) refers to the extent to which a person desires that one's in-group dominate and be superior to out-groups. SDO is considered to measure social and political attitude orientation toward inter-group relations, reflecting whether one generally prefers such relations to be equal, versus hierarchical, that is, ordered along a superior-inferior dimension (Pratto et al. 1994). We used the updated SDO7 (Ho et al. 2015) which is significantly shorter (8 items) and uses more modern references compared to original iterations. Both the RWA and the SDO scales are said to capture two different dimensions of political opinions (Asbrock et al. 2010). This two-dimensional interpretation offers a substantial amount more nuance to our interpretation of Colombian voters. Scales can be seen as capturing different aspects of political values, goals, and motivations. Nonetheless, the scales should show a reasonable degree of correlation in capturing left/right political view points. Both scales were measured on a 5-point Likert-like scale (original scales used a 1–9 rating). However, for ease of use with mobile phones we shortened the scale for compatibility. Importantly, all political scales were measured on the same scale to allow for equivalence. Higher ratings on both the SDO and the RWA suggests a stronger right-wing political ideology.

The *Colombian specific political questionnaire* was developed during the piloting section of the experiment through participant suggestions for extensions of either the RWA or SDO scales. Although both scales have been used in cross-cultural contexts (Duckitt et al. 2010), including other collectivist cultures in South America (Cantal et al. 2015), neither has been validated in a Colombian setting to our knowledge. Specifically, candidates' proposals addressed the Colombian political climate: political nuances are idiosyncratic to the electoral timeframe. We concluded that such a measure would be highly beneficial in assessing Colombian political dimensions and to counter limitations in the RWA and SDO. We chose 9 items to represent several key components of the current elections in Colombia – three proposals by each candidate that identified the political "center", "right", and "left" were rated by each participant, also on a 5-point scale (see full scale in the complementary website[¶]).

An additional measure of the Big-Five personality traits was initially considered given the link between personality traits and political ideology. However, to limit the time commitment required by participants in the experiment this was not included.

Music selection

We refer the reader to studies by Quishpe (2020), Barbosa Caro and Suavita (2019), and Katz-Rosene (2017) with respect to historical, functional, and lyrical analysis of the two types of music used: (1) FARC-songs (mainly in the style of *vallenato* and *canción social*) and (2) AUC-songs (in the style of *corridos*). These musical styles make part of traditional Colombian (and Latin-american) music, yet they have distinctive sonorities, structures, and instrumentation. It must be noted that music with politically motivated lyrics from both types have incorporated other similar styles of music as well (e.g., hip-hop and rock), but this study only considers this reduced range of styles. Additionally, FARC-songs have been typically created by active members from the guerrilla as a mechanism of identity confirmation and propaganda (Quishpe 2020), while AUC-songs have been typically been produced by sympathizers of the paramilitaries as promotion to their deeds and in open criticism to the FARC and left-wing politicians (Barbosa Caro and Suavita 2019). Crucially, the functionality of the music and the target listener can be seen as different.

We remark that humans frequently listen to music *without* feeling any emotion at all (Kivy 1990; Juslin 2019), but music *might* trigger mechanisms such as episodic memories for particular individuals (Juslin 2013; Eerola 2018). However, the potential induction of emotions from the music in this study is based mainly to the semantic content of the lyrics – inducing different emotions to listeners with different political views. In Gómez-Cañón et al. (2021), we exclusively evaluated acoustic features from the music – in this study we extended the analysis to features in the lyrics as analyzed in natural language processing and topic modeling (see the personalization section for the description of computational models). Nonetheless, the acoustic features are useful to provide a content-based contrast among the different styles of music: (1) FARC-songs typically use less instruments and might include only voice and guitar, and (2) AUC-songs are more heavily orchestrated with faster tempo. Namely, the machine learning models should be able

[¶] https://juansgomez87.github.io/2022/06/musicandscience

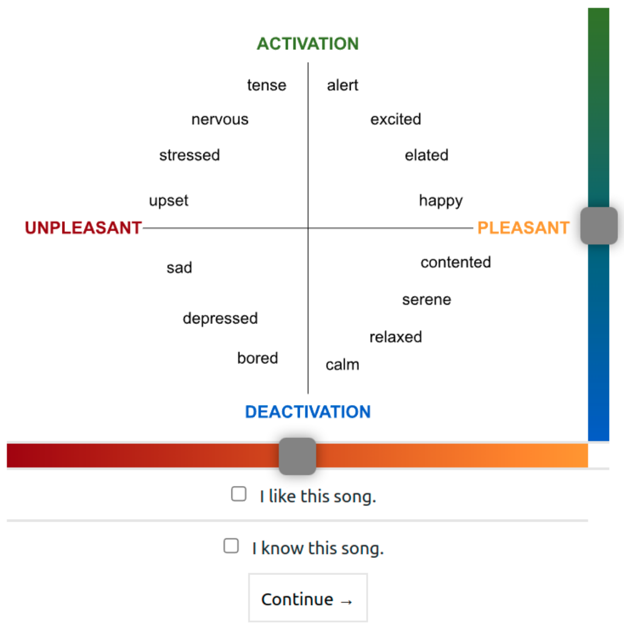

Figure 1. Annotation interface for the experiment.

to differentiate between the types of music – the interesting element is to attempt to understand which users will provide problematizing labels (i.e., music that induces subjectively negative emotions) that can bias the algorithm towards a particular class.

We used 50 music excerpts with lyrics from each music type (30 seconds long) and extracted 260 emotionally relevant acoustic features (mean and standard deviation of 65 low-level music descriptors and their first order derivatives) from segments of 1 second (Aljanaki et al. 2017), with 50% overlap, and standardize across features – using the IS13 ComParE feature set (Weninger et al. 2013) and OpenSMILE toolbox (Eyben et al. 2013). We processed the tracks using Audiosourcere DeMIX software to extract versions without lyrics – we obtained a total set of 150 excerpts (50 FARC-songs, 50 AUC-songs, 25 FARC-songs without lyrics, and 25 AUC-songs without lyrics). Each excerpt was normalized for loudness following the ITU-R BS.1770-4 recommendation using the pyloudnorm package[‖].

Annotation gathering

We use a discretized model of emotion based on Russell's circumplex model (Russell 1980) and recent work on MER (Panda et al. 2018; Gómez-Cañón et al. 2021a), which conceptualizes emotions in to two-dimensional core affect (arousal and valence) and four distinct categories/quadrants of emotion: Q_1 (positive valence and arousal), Q_2 (positive arousal and negative valence), Q_3 (negative valence and arousal), Q_4 (negative arousal and positive valence). Figure 1 shows the annotation interface: Q_1 refers to emotions such as happy and excited, alert; Q_2 refers to emotions as tension and anger; Q_3 refers to emotions as sadness and boredom; Q_4 refers to emotions as calmness, serenity. To refer to arousal, we used the words activation/deactivation (*activación/desactivación*). To refer to valence, we used pleasant/unpleasant (*positivo/negativo*). Annotations of arousal and valence were made on continuous sliders ranging from 1–100. We use continuous scale values to analyze annotations but use the discretized classes to train our machine learning models. We also collected each participant's preference and familiarity for the musical excerpts through check boxes "*I know this song*" and "*I like this song*" (see figure 1).

Personalization

We use the "machine consensus" MER personalization strategy presented by Gómez-Cañón et al. (2021b): consensus entropy for active learning. This strategy uses a committee of classifiers to analyze their output agreement and queries each user for instances with the highest uncertainty. Each participant receives a committee of classifiers: 5 independent extreme gradient boosting models (Chen and Guestrin 2016) and 5 logistic regression models optimized with stochastic gradient descent (Bottou 2010). Each one of these models had previously been pre-trained on separate cross-validation splits of the DEAM dataset, the benchmark dataset for MER (Aljanaki et al. 2017). In order to select uncertain data to be labeled, classifiers predict the output probabilities for the pool of excerpts. We then perform the consensus entropy strategy by analyzing the disagreement across classifiers. For example, full disagreement from a committee of four classifiers results when each one predicts a different class/quadrant with 100% probability. This yields average probabilities per quadrant $p_{avg} = \{Q_1 : 0.25, Q_2 : 0.25, Q_3 : 0.25, Q_4 : 0.25\}$ and high inter-class entropy/uncertainty of 1.386. Following Gómez-Cañón et al. (2022), we balanced the instances with respect to the quadrants for each epoch: (1) prior to the calculation of entropy, we split the probabilities p_{avg} into four matrices corresponding to the instances with higher probability of belonging to each quadrant, (2) we calculate entropy independently for each matrix (four quadrant probabilities $\times$ 150 instances), and (3) we select instances with highest entropy from each matrix. Thus, we alleviated the issue of imbalanced classes for each retraining iteration, since the instances selected for query are more likely to belong to each of the quadrants. In the case that the probabilities do not favor a particular quadrant (i.e., models are biased towards particular classes), we simply select the instances with highest entropy from the initial matrix. Excerpts with highest uncertainty are then queried to each participant to be annotated. Initially, we pseudo-randomly draw 2 excerpts from each type of music (6 excerpts for the first annotation iteration), retrain our classifiers with the annotations provided by each user, identify the excerpts to be annotated for the next iteration, and present the new batch of music to be annotated. Given the low amount of available music, we perform only five iterations for a total of 30 annotations per user – past research has shown that only 20-30 annotations are needed in order to reach personalization (Su and Fung 2012; Chen et al. 2017). Please refer to Gómez-Cañón et al. (2021b) and Gómez-Cañón et al. (2022) for additional information of the consensus entropy methodology.

[‖] https://github.com/csteinmetz1/pyloudnorm

TM approach	Algorithm	5 topics	10 topics	15 topics	20 topics	25 topics	30 topics	35 topics
	NMF	0.556	0.588	0.464	0.551	0.628	0.625	0.551
Classical TM	**SVD**	**0.711**	**0.687**	**0.714**	**0.709**	**0.723****	**0.675**	**0.777**
	LDA	0.527	0.428	0.497	0.518	0.469	0.641	0.516
	GS-DMM	0.622	0.642	0.650	0.595	0.610	0.587	0.556
Short text TM	NQTM	0.597	0.557	0.549	0.526	0.491	0.528	0.557
	Biterm	0.431	0.569	0.379	0.544	0.515	0.453	0.524

Table 1. Classical and short text topic modeling (TM) approaches tested to produce lyrics models. Bold indicates the best algorithm and ** indicates the amount of topics selected (100 songs with lyrics × 25 topics).

In order to extract information from lyrics, we used a standard topic modeling approach (Manning and Schütze 1999) – an unsupervised method that detects word patterns within different texts and attempts to cluster documents (i.e., lyrics) into a particular amount of topics. The classical bag-of-words approach was implemented by: (1) calculating the frequency of words from each lyric (i.e., term frequency-inverse document frequency); (2) testing different algorithms to obtain a numeric representation of the likelihood of each lyric to belong to topic t; (3) using the extracted text features as input to a logistic regression classifier that is subsequently trained with the annotations of each participant. However, short texts face the challenge of being ambiguous and noisy for topic modeling (Albalawi et al. 2020) – we use the text of the lyrics from the 30 seconds selection exclusively. Following Valero et al. (2022), we tested different classical and short text topic modeling methods, and evaluated binary classification (FARC-songs or AUC-songs). Thus, in (2) we tested non-negative matrix factorization (NMF), singular value decomposition (SVD), latent dirichlet allocation (LDA), collapsed Gibbs sampling for dirichlet multinomial mixture (GSDMM - Yin and Wang (2014)), negative sampling and quantization topic model (NQTM - Wu et al. (2020)), and biterm topic model (Biterm - Yan et al. (2013)). We performed 5-fold cross-validation and report F1-scores in table 1. In essence, we varied the amount of topics and selected the algorithm that offered best classification performance and topic coherence – singular value decomposition and 25 topics. We obtained a feature matrix that represents the data from the lyrics and we use as input to the classifier: 100 songs with lyrics × 25 topics calculated using SVD.

In summary, we produce four types of models: *ACO* models use acoustic features, *LYR* models use features extracted using topic modeling on the lyrics, *MIX* models with use both acoustic and lyrics features, and *RAND* models that pseudo-randomly present music to be annotated (i.e., no entropy is calculated).

Results

A total of 194 participants started the experiment. 52 participants completed the whole study. 3 participants were removed from the completed entries since the server failed to collect their annotations, leaving a total of 49 participants for the analysis. We gathered participants from different age groups ($\mu = 35.6, \sigma = 12.75$): 19 participants were 18-30 years old, 23 participants were 30-50 years old, and 7 participants were 50-65 years old. We had unbalanced participation with respect to gender: 19 female, 29 male, and 1 non-binary. With respect to musical self-identification, most of our participants were non-musicians**. Moreover, 15 participants received the *MIX* model, 13 received the *RAND* model, 11 received the *LYR* model, and 10 received the *ACO* model.

Political scale assessment

We obtained five political scores s ranging from 1 to 5 from our three questionnaires: RWA (s_{RWA}), SDO (s_{SDO}), PlanLeft (s_{left}), PlanCenter (s_{center}), and PlanRight (s_{right}). Agreement to statements from the Colombian specific political questionnaire corresponded to agreeing with the political discourse from a given candidate (i.e., s_{left}, s_{center}, and s_{right}). First, the five political ideology scales were correlated to assess their effectiveness at capturing political leaning. All five scales correlated as predicted showing that the RWA and the SDO scores capture left/right leaning ideas in the Colombian population, presented the our Colombian specific political questionnaire (see figure 2).

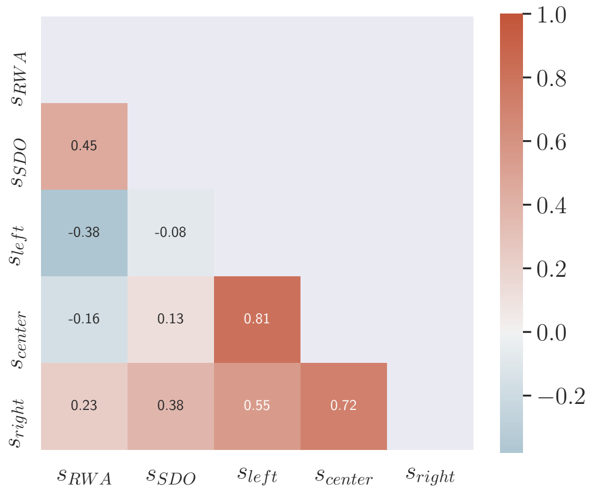

Figure 2. Correlation matrix for political scales measured in the experiment.

**27 non-musicians, 6 amateur musicians, 5 serious amateur musicians, 5 professional musicians, 3 semi-professional musicians, and 3 music loving musicians

		All users			Left (n=15)			Center (n=22)			Right (n=12)			Rand (n=13)		
		Q	A	V	Q	A	V	Q	A	V	Q	A	V	Q	A	V
All	α_k	0.035	0.025	0.051	-0.047	-0.088	-0.067	0.045	0.018	0.131	0.028	0.058	-0.013	0.058	-0.036	0.181
	α_c	0.921	0.917	0.937	0.863	0.884	0.947	0.872	0.852	0.894	0.853	0.823	0.819	0.885	0.867	0.938
Lyr.	α_k	0.028	0.025	0.019	-0.033	-0.098	-0.054	0.045	0.040	0.072	0.031	0.032	-0.006	0.053	-0.032	0.158
	α_c	0.876	0.865	0.915	0.700	0.800	0.920	0.803	0.820	0.852	0.785	0.779	0.836	0.800	0.769	0.920
No Lyr.	α_k	0.052	0.027	0.122	-0.082	-0.082	-0.093	0.047	-0.033	0.260	0.024	0.133	-0.031	0.069	-0.037	0.230
	α_c	0.868	0.808	0.849	0.767	0.705	0.893	0.730	0.510	0.757	0.735	0.310	0.223	0.774	0.731	0.771

Table 2. Inter-rater reliability and consistency statistics. We report Krippendorff's α_k and Cronbach's α_c. Q stands for quadrants, A for arousal, and V for valence.

Next, we segmented the participants into three groups: "right-leaning", "left-leaning", and "center" orientation. In order to produce these groups: (1) we grouped participants using the 33% quantile from the RWA and the SDO scores –namely, $s < 0.33$ are "left", $0.33 \leq s \leq 0.66$ are "center", and $s > 0.66$ are "right"; (2) we group participants from the Colombian specific questionnaire depending on the highest score obtained – for example, we assign "left" if $s_{left} > s_{right}$ and $s_{left} > s_{center}$, and (3) we obtain the final group by taking the mode from the three resulting classes – for example, if RWA results in "right", SDO results in "center", and the Colombian specific questionnaire results in "right", we assign the participant to the group "right". Using this score, our participants were grouped as follows: 22 participants are "center", 15 participants are "left-leaning", and 12 participants are "right-leaning".

Annotation analysis

As we have argued previously in Gómez-Cañón et al. (2021a), inter-rater agreement must be routinely analyzed and reported in studies that involve MER. In table 2, we summarize inter-rater reliability statistics calculated from the discretized categories used for the MER algorithms: nominal Krippendorff's coefficient α_k and Cronbach's consistency coefficient α_c. To calculate the statistics, we keep only the songs that have been annotated by at least two participants (140 songs from 150 songs) – this reveals that given the initial seed of 6 songs, the response and model variability allowed the participants to annotate most of the data. As mentioned previously, the *RAND* model would produce pseudo-random presentations of songs to be annotated – using the same random seed resulted in 13 participants that annotated the same 30 songs (9 AUC-songs, 12 FARC-songs, and 9 songs without lyrics). Moreover, all participants annotated the same initial 6 songs. We discuss inter-rater reliability and consistency statistics as follows: (1) agreement as measured by α_k is notably low in general – we argue that the sparsity of the annotations leads to increasing the probability of agreement due to chance and lowering the agreement coefficient (i.e., each participant annotated 30 from a pool of 150 songs); (2) annotations of valence are more consistent than annotations of quadrants and arousal for most groups (see α_c) – this is a surprising finding, since typically valence is the most subjective quality and exhibits least consistency; (3) annotations of music without lyrics show a low consistency for "right-leaning" participants – it is likely that the music without lyrics is interpreted with more freedom (e.g., vallenatos and corridos are music normally used for parties); (4) as expected, the response variability of induced emotions from music is evident – using personalized models to capture response variability is a reasonable approach to create MER models (Yang et al. 2007; Gómez-Cañón et al. 2022).

Linear Mixed Model (LMM) analysis was carried out separately for continuous rating of arousal and valence with five fixed factors (model [*RAND*, *LYR*, *MIX*, *ACO*], type of music [FARC-songs/AUC-songs], political orientation of the participant [Left/Center/Right], lyrics [LYRICS/No LYRICS], and gender), and two random factors (participant and track). For valence, main effects for the *MIX* model ($\beta = -31.4$, *t*=-2.23, *p*=.031) and gender ($\beta = 23.5$, *t*=2.33, *p*=.025) emerged. We refer the reader to figure 3 for a visualization of the annotations of music with lyrics. Planned contrasts in valence annotations showed significant differences in AUC and FARC music only in the *LYR* model for right-leaning participants annotations of music without lyrics ($\beta = -62.8$, *t*(1352.0), *p*=.033). For arousal, main effect for gender ($\beta = -15.0$, *t*=-2.03, *p*=.049) and an interaction between type of music and political orientation [Right] ($\beta = -17.4$, *t*=-2.63, *p*=.009) emerged. Planned contrasts showed that arousal annotations were significantly higher for AUC songs compared for FARC songs for both right-leaning participants in both music with lyrics (*t(1348.6)*=3.31, *p*<=.001) and without lyrics (*t(1244.4)*=2.00, *p*=.046). Similarly, center-leaning participants showed significantly higher for AUC songs compared for FARC songs for music with lyrics (*t(1095.5)*=2.55, *p*=.011). Differences in right-leaning participants were more pronounced with annotations obtained under *MIX* (β=17.7, *t*(1355.0), *p*=.081) and *LYR* models (β=29.2, *t*(1354.1), *p*=.002). Differences in center-leaning participants were more pronounced only in the *LYR* model (β=19.1, *t*(1357.3), *p*=.022). No differences emerged for left-leaning participants (β=8.7, *t*(1313.5), *p*=.11) or between participant's political orientation and the type of tracks under the random model and a model based on acoustic features ($p >$.20).

Algorithmic evaluation

In order to evaluate the personalized models, it must be noted that there was *no testing data* – all the annotated data was used to train the models and there is no "ground truth" to compare the predictions from the models. However, we propose an evaluation strategy to account for the possible bias that results from "misusing" the personalized models. Given that there is a significant difference in the way that participants from different groups annotated the music, it is likely that models are inherently biased towards certain

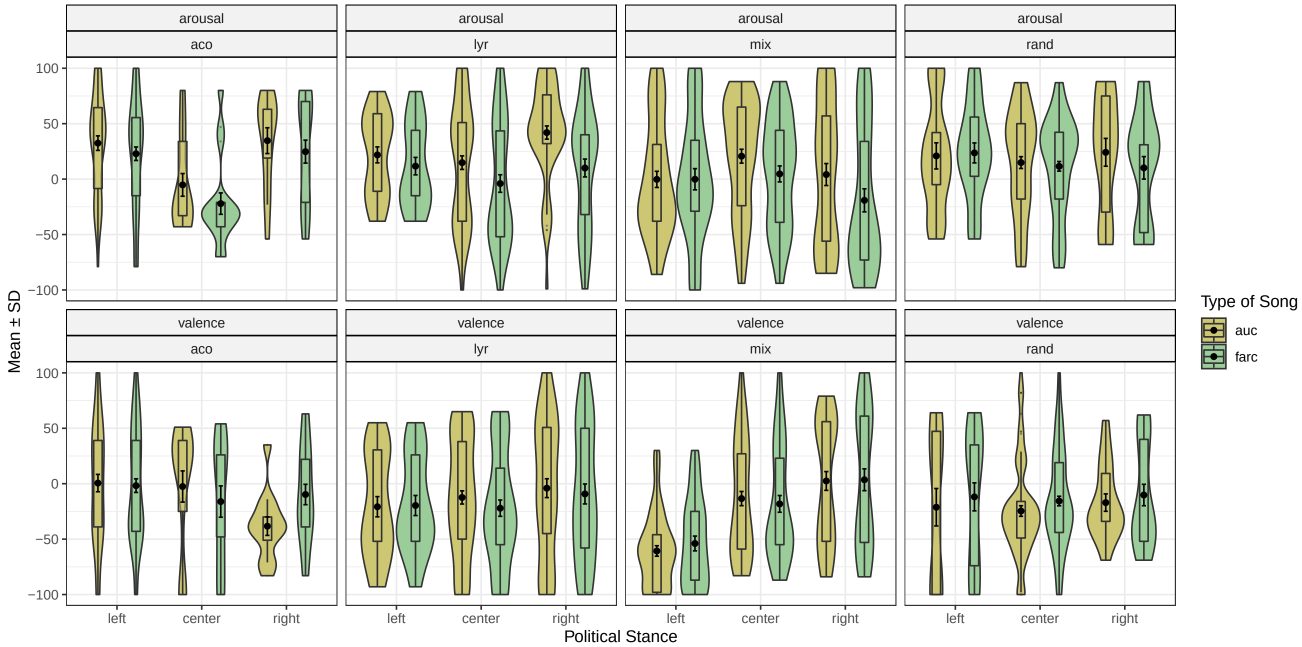

Figure 3. Annotation analysis of music with lyrics showing the comparisons with respect to model type, annotation (continuous arousal and valence), and political stance.

Political leaning	Model type	Top 10 AUC (%)	Top 10 FARC (%)	Top 10 No Lyr. (%)	Top 20 AUC (%)	Top 20 FARC (%)	Top 20 No Lyr. (%)
Center	Acoustic	35.0	65.0	0.0	40.0	60.0	0.0
	Lyrics	42.0	50.0	8.0	48.0	42.0	10.0
	Mix	53.3	35.0	11.7	45.0	40.8	14.2
	Random	34.4	53.3	12.2	38.3	48.9	12.8
Left	Acoustic	43.3	**51.7**	5.0	43.3	**46.7**	10.0
	Lyrics	**65.0**	35.0	0.0	**55.0**	45.0	0.0
	Mix	**50.0**	30.0	20.0	**50.0**	29.0	21.0
	Random	**35.0**	**35.0**	30.0	35.0	27.5	**37.5**
Right	Acoustic	**50.0**	40.0	10.0	42.5	**45.0**	12.5
	Lyrics	**45.0**	22.5	32.5	**38.8**	28.8	32.5
	Mix	25.0	**40.0**	35.0	25.0	**38.8**	36.3
	Random	30.0	**45.0**	25.0	22.5	**40.0**	37.5

Table 3. Proportion of the top 10 and top 20 aggregated predictions from personalized models (in this case, for predictions of negative valence). Bold indicates highest proportion between AUC, FARC and songs without lyrics.

categories (Nigam et al. 2000; Schölkopf et al. 2007) – moreover, the only evaluation possible is probabilistic given the choice of the algorithms for machine learning. Thus, each personalized model was used to test on the remaining data from each participant, following our previous work on Gómez-Cañón et al. (2021) – the resulting matrix has 120 songs $\times$ 4 quadrants. We sorted the matrix with respect to the highest probabilities of belonging to a particular class and select the top 10 and 20 predictions that: belong to a particular quadrant (i.e., Q_1, Q_2, Q_3, and Q_4), belong category of arousal (i.e., positive and negative), and belong to a category of valence (i.e., positive and negative). Table 3 summarizes the findings that we deem ethically problematic – the type of music that some personalized models appear to classify with high probability of negative valence is revealing the political stance from the participants. In general, we conclude that results are not as discernible as the ones reported in Gómez-Cañón et al. (2021): (1) in the case of "left-leaning" participants, the *LYR*, *MIX*, and *RAND* models predict that AUC-music (i.e., "right-wing") has a higher probability of inducing negative valence than FARC-music; (2) in the case of "right-leaning" participants, we find that only the *MIX* and *RAND* models predict that FARC music (i.e., "left-wing") has a higher probability of inducing negative valence than AUC-music; (3) we find no trends as to the type of models that might capture the political stance (i.e., only the *MIX* models do so consistently for the left- and right-leaning groups) – it is likely that the classification strategy might be too coarse to capture the response diversity of each participant and subtle political differences; (5) despite that the political stance from the participants was not necessarily captured by all the models, we find that the models are accurately capturing that *both* types of music might induce emotions with negative valence – this was expected since the political content of the lyrics was strong and specific memories with a negative connotation might have been triggered through music listening; (6) we find that certain participants resulted with models that are the most problematic, we offer figures of predictions in the complementary website for clarity on the behavior of each participant.

Discussion

We discuss our findings regarding our proposed hypothesis as follows:

H1 Participants with different political viewpoints (goal-relevance) will show different induced arousal annotations.
H2 Politically sensitive music stimuli that agree/disagree with participants political stance (goal-congruence) will show different induced valence annotations.
H3 Personalized algorithms will be effectively biased towards specific categories depending on the political stance of listeners – namely, personalized algorithms might reflect the listeners' political views.

In relation to our first hypothesis H1, these results showed an overall effect for arousal annotations. Planned contrasts indicate that right-leaning participants arousal annotations of FARC and AUC music differed in music with and without lyrics. This effect is most present in *LYR* and *MIX* models for music with lyrics, and in the *MIX* model for music without lyrics. Gender also played a significant role. When visualizing the results, figure 3 shows that "right-leaning" participants produced significantly higher arousal annotations for AUC music (politically-right aligned) over FARC music (politically-left). Moreover, the difference between arousal annotations in different MER models suggests that the *LYR* models amplify these effects leading to higher annotations in arousal. Although the overall trend for higher arousal annotations for AUC music is broadly preserved across models this provides support for our assumption that lyrics (at least in combination with acoustic features) plays a key role in induced emotion. All together these results provides partial support for H1, that goals (as political social identities) influence participants arousal annotations, though this trend was not observed in left-leaning participants.

In relation to our second hypothesis H2, overall effects for valence ratings were noted in the MIX model and for gender. Planned contrasts showed one significant difference between AUC and FARC annotations for right-leaning participants in the *LYR* models for music with no lyrics. These findings do not support H2, that congruence between music type and political ideology would influence annotations of valence positively and vice-versa. Instead what was observed was significantly lower valence annotations for "left-leaning" participants for both AUC and FARC music, in comparison to "center" and "right-leaning" participants. This instead suggests large individual differences in valence ratings are acknowledged in the interaction between acoustic qualities and lyrics. Although the specific hypothesis is not supported the results do support the overall aim of the study: to show that political ideologies do influence ratings of valence when acoustic and lyrics are considered.

Interpreting these results through a goal-directed mechanism suggests that participants with a "right-leaning" political ideology may have found politically-right AUC-music more engaging than FARC-music. This effect is shown beyond a purely acoustic interpretation of the arousal annotations. That is, AUC music is typically instrumentally more rich and faster in tempo. These acoustic features are often linked to higher arousal ratings (Céspedes-Guevara and Eerola 2018). However, the same effects were not observed in music without lyrics, as would be expected in a solely acoustic interpretation. A goal-directed interpretation would equate these difference to greater attention and cognitive resources being focused upon the stimulus; a process that would subsequently lead to higher subjective arousal annotations.

One possibility not explored here is that results suggest an interaction between valence and arousal annotations which has not been assessed. This form of analysis allows us to look at specific predictions in annotations of core-affect but does not allow us to identify more global changes in core-affect. A combined analysis would be more inline with theoretic descriptions of core-affect as a single entity, the two dimensions inseparable at higher levels of cognitive processing (Russell 2003). Beyond our hypotheses, we show lyrics to be an important factor in producing these differences in core-affect. This further supports a goal-directed interpretation over a purely musical one – lyrics more readily orientating awareness of a stimulus as related to an individuals social identity. Yet, some degree of associative processing of the music as related to social-identity must also be true as the effects were also seen in right-leaning participants annotations of non-lyrical music too. Not in relation to the hypotheses, a significant effect was also observed for center-leaning participants in the *LYR* model for music with lyrics. This was not expected but adds greater support to the idea that the lyrics are important in induced musical emotions, at least pertaining to arousal. However, these effects may be an artifact of participants political groupings – a center-right political ideology that is not clearly distinguished in this participant grouping.

In relation to our third hypothesis H3, we find conflicting results to our preliminary findings in Gómez-Cañón et al. (2021). To build upon our previous study, in this experiment we added several more levels of complexity: we used music without lyrics in the same styles, we included the *LYR*, *MIX*, and *RAND* models, and we allowed our participants to annotate the music using continuous arousal and valence sliders. We find particularly interesting that all models were able to capture that music belonging to both music styles (which is a strong emotional stimuli) would induce negative valence. Table 3 shows that most of the models for all groups of participants would predict that music with lyrics would induce negative valence, as opposed to music without lyrics. Thus, we argue that the algorithms have effectively captured that music with negative valence belongs primarily to both AUC- and FARC-music through the personalization strategy. To a certain extent, the machine learning algorithms are able to capture the political stance of some participants (i.e., the model from a "left-leaning" participant would show high probability that AUC-music ("right-wing") would induce negative valence). However, the political views of the personalized algorithms were not necessarily reflected from each one of the models. In general, the broad assumption behind this experiment was that a strong political stances should be reflected accordingly with arousal and valence annotations – however, it is likely that participants from different political views would find both types of music to produce emotions with negative valence similarly. We do not find strong support that the models have effectively captured the political stance from the participants. Yet, it is important to remark that algorithms were effectively capturing that both types of music would induce a negative emotion (see table 3).

Beyond the initial hypotheses we set out, we have also provided tentative evidence for the validity of the right-wing authoritarian (RWA) scale and the social dominance orientation (SDO) scale in a Colombian population. This is tentative because the sample size is too small to draw conclusions about the population generally. Yet, they behave and correlate as expected with one another and with our Colombian specific political questionnaire. The interpretations of the SDO and the RWA is highly important to our interpretation of the data. A critical look at these scale may shed light on why the effects were explicitly seen in "right-leaning" participants. In relation to H1 and H2 and the valence and arousal annotations, many of the observed effects were noted specifically in "right-leaning" participants. However, the distribution of participants in the SDO and the RWA both produce a left-skew suggesting greater "leaf-leaning" ideologies (supported by the results of the first round of presidential elections). This may mean that some effects in the "left-leaning" participants are somewhat muted by the "center" grouping.

Limitations and future directions

We note several limitations in the design of the experiment and some future directions for researchers in music cognition and MER to take.

- L1 Sample size: the sample size was small and this increases the chances of a type II error occurring. Regardless, the hypothesized effects were observed in the data. It is, however, difficult to draw conclusions about the size of these effects. Similarly, the comparisons with many factors (e.g., between model types) may not be statistically strong enough to allow smaller effects to be observed.
- L2 Dropout rate: The number of participants who started the experiment but did not complete it was substantially higher that is typical for online experiments (Hoerger 2010). This may have been due to the length of the experiment as Hoerger suggests. It may also have been due to the nature of experiment. That is, attempting to allow algorithms to identify music that can induce negative effect in a participant. If the algorithms were even somewhat successful in this task, it is quite possible that people did not find the experiment enjoyable and left. More data could have been drawn from the numerous partial responses if participants had completed the political scales. However, if the political scales had been placed earlier in the experiment design, it may have cued in participants to the main manipulation in the experiment and skew results.
- L3 Personality traits: we have cited lots of literature linking personality traits with political leaning. Moreover, as noted in the introduction section, music preferences similarly have studies linking them to personality traits and to political ideologies. MER has similarly began to acknowledge the benefit of collecting such variables for predicting music preference (Zangerle et al. 2021) and understanding physiological responses to music (Hu et al. 2018). This study did not have the space in the experimental design to include these variables, though we hypothesize that this would be a valuable and informative direction for future studies to take, in either in MER or affective music cognition.
- L4 Type of machine learning models: the type of algorithms that were used for the personalization approach have been established as classical approaches to classification and are efficient models. However, the logistic regression classifier (used in the *ACO*, *LYR*, and *MIX* models) assumes linearity between the acoustic or lyrics features and the annotations from participants – it is likely too coarse to model the subtle non-linearity that relate features to an annotation. Our previous work in Gómez-Cañón et al. (2021b) shows that faster personalization could be achieved by the use of convolutional neural network architectures – given the computational requirements of deep learning model estimations, our web servers were not capable of supporting online training.
- L5 Cultural specificity: The study has limitations in its generalizability to other countries, populations, or political spheres. Colombia is quite unique in that it has such specific music genres that culturally associate with particular political movements. How this would be represented in other countries or to what degree music could be said to specifically relate to political ideologies would take careful consideration of the literature to replicate.
- L6 Political labels: The measures for political leaning produced more a predominant left skew in both the RWA and SDO. This in itself did not drastically affect the analysis; three almost perfectly even groups were still derived from our grouping through percentiles. With regards to the interpretation of the results. However, it is fair to acknowledge that "right-leaning" participants may be better described as "center-right" to "right" political ideologies. "Left-leaning" participants cause similar problems as it is difficult to find a clear distinction between "left" and "center-left" ideologies.
- L7 Musical extremes: It is of course true that these music genres represent quite extreme political ideologies within the Colombian community. We in contrast, have used them as representative of right/left political distinctions. It is quite possible that they are more representative of far-right/left political ideologies.
- L8 RWA and SDO scales: The Likert-like scales used to measure these two political attitudes were shortened from their original range of 1–9 to 1–5 to allow for compatibility and ease of use with the online platform. This raises possible questions about reliability and

validity of the scales in comparison to the original development. In relation to the current study this may be most apparent in the scales' ability to discriminate between political ideologies. This may be particularly evident when acknowledging the left skew observed in this experiment in both scales, suggesting a possible loss of clarity between center and left participants.

Music, persuasion, and... manipulation?

It has been argued that music and persuasion have indirect relationships and are never strictly causal – as the persuasion/manipulation of a person is already a difficult task, it can only be "helped" or "promoted" by music. As mentioned by Herrera (2009), music can only contribute as a "persuading factor" to an induced emotional state, which can be associated between music and a particular person or message, contributing to reevaluating attitudes and actions. Indeed, studies regarding the impact of musically-induced emotions on decision making have been studied since 2006 – as reviewed by Palazzi et al. (Palazzi et al. 2019). While the diversity of theories regarding the underlying mechanisms of emotion induction with music is still debated (Damasio 1994; Gabrielsson 2006; Hansen and Christensen 2007), it has been argued that music is a powerful and engaging stimulus that can promote prosociality (Ruth 2018; Ruth and Schramm 2021), impact customer behaviors (Hansen and Christensen 2007), and influence processes of decision-making and risk-aversion (Fischer and Greitemeyer 2006; Greitemeyer 2011).

In essence, we would like to propose a naive reflection from the findings of this experiment. While research in music to promote well-being and beneficial uses has promoted important research and is likely to grow in the following years (Hu et al. 2021; Agres et al. 2021), we stress that research in music-induced harm has only started to be explored (Saarikallio et al. 2015; Sharman and Dingle 2015; Silverman et al. 2020; Alluri 2020). It is critical that the field of music technology acknowledges and builds upon the field of music psychology – there is a necessity to ground technological applications on reliable psychological research, since each algorithm will be used to develop specific use cases. Moreover, there is a need to acknowledge that technology poses asymmetrical power relationships onto vulnerable populations (Mohamed et al. 2020; Adams 2021) – while the episode from Cambridge Analytica is a well-known situation in Western societies, it is less-known that they influenced elections in more than 30 countries and 100 election campaigns (including Colombian elections). To the researchers in music technology (and mainly the Music Emotion Recognition task), we respectfully suggest: before engaging with implementing the latest machine learning algorithm, assembling enormous datasets, or getting state-of-the-art accuracy, we believe that the most relevant question that should be addressed is *what for?*

Conclusions

In relation to the broader research questions posed in our study:

R1 Do an individuals' political values (goals) influence emotional episode induced by music?

R2 Can a MER algorithm for induced emotions be biased towards a particular opinion with respect to music with polarizing lyrics?

This study supports the overall aims of the study, showing that an individual's political-identity makes a meaningful contribution to their induced emotional experience of music, at least rated in terms of core-affect. These effects were observed beyond an entirely acoustic interpretation. Yet, how this relates explicitly to a goal-directed interpretation and the extent to which goal-directed mechanisms can influence emotions induced by music remains an open question. Moreover, this effect can be manipulated by algorithmic models to bias individuals towards negative emotional states. We find this to be the most significant finding from this research, and should be evaluated toward understanding how MER algorithms with increasing personal data could be used, both in positive and negative scenarios.

Acknowledgements

We would like to thank Alastair Porter (MTG) for his invaluable help and patience when staging the online platform. Also to Elena Epure (Deezer) for her guidance on creating the lyrics classifications algorithms.

The research work by JSGC is partially supported by the European Commission under the TROMPA project (H2020 770376) and the Project Musical AI - PID2019-111403GB-I00/AEI/10.13039/501100011033 funded by the Spanish Ministerio de Ciencia, Innovación y Universidades (MCIU) and the Agencia Estatal de Investigación (AEI). Additional funding came from the Arts and Humanities Research Council scholarship (grant code: AH/L503927/1) awarded to TML through the Northern Bridge Doctoral Training Partnership.

References

Adams R (2021) Can artificial intelligence be decolonized? *Interdisciplinary Science Reviews* 46(1-2): 176–197. DOI: 10.1080/03080188.2020.1840225.

Agrawal Y, Shanker RGR and Alluri V (2021) Transformer-based approach towards music emotion recognition from lyrics. In: Hiemstra D, Moens MF, Mothe J, Perego R, Potthast M and Sebastiani F (eds.) *Advances in Information Retrieval*. Springer International Publishing, pp. 167–175.

Agres KR, Schaefer RS, Volk A et al. (2021) Music, Computing, and Health: A Roadmap for the Current and Future Roles of Music Technology for Health Care and Well-Being. *Music & Science* 4: 1–32.

Albalawi R, Yeap TH and Benyoucef M (2020) Using Topic Modeling Methods for Short-Text Data: A Comparative Analysis. *Frontiers in Artificial Intelligence* 3. DOI:10.3389/frai.2020.00042.

Aljanaki A, Yang YH and Soleymani M (2017) Developing a benchmark for emotional analysis of music. *PLoS One* : 1–22DOI:10.1371/journal.pone.0173392.

Alluri ASYGMSSHSV (2020) Tag2Risk: Harnessing Social Music Tags for Characterizing Depression Risk. In: *Proceedings of the 21st International Society for Music Information Retrieval Conference*. Montréal, Canada, pp. 384–391.

Altemeyer B (1998) The Other “Authoritarian Personality”. *Advances in Experimental Social Psychology* 30: 47–92.

Arocha R J et al. (1988) *Colombia: Violencia y Democracia*. Universidad Nacional de Colombia.

Asbrock F, Sibley CG and Duckitt J (2010) Right-wing authoritarianism and social dominance orientation and the dimensions of generalized prejudice: A longitudinal test. *European Journal of Personality* 24(4): 324–340. DOI:10.1002/per.746.

Asutay E and Västfjäll D (2021) The goal-relevance of affective stimuli is dynamically represented in affective experience. *Royal Society open science* 8(11): 211548.

Balleine B and Dickinson A (1998) *Consciousness—The interface between affect and cognition*, chapter 4. Oxford University Press, pp. 57–85.

Barbosa Caro E and Suavita JR (2019) Paramilitarism and music in Colombia. *Journal of Language and Politics* 18(4): 541–559. DOI:10.1075/jlp.19019.bar.

Barrett LF (2006) Solving the emotion paradox: Categorization and the experience of emotion. *Personality and social psychology review* 10(1): 20–46.

Barrett LF (2017) *How emotions are made: The secret life of the brain*. Houghton Mifflin Harcourt.

Barthet M, Fazekas G and Sandler M (2013) Music Emotion Recognition: From Content- to Context-Based Models. In: *From Sounds to Music and Emotions*. Berlin, Heidelberg: Springer, pp. 228–252.

Berquist CW (1978) *Coffee and Conflict in Colombia, 1886-1910*. Duke University Press.

Birhane A and Guest O (2021) Towards Decolonising Computational Sciences. *Kvinder, Køn & Forskning* 29(1): 60–73. DOI:10.7146/kkf.v29i2.124899.

Bottou L (2010) Large-Scale Machine Learning with Stochastic Gradient Descent. In: *Proceedings of the 19th International Conference on Computational Statistics (COMPSTAT)*. pp. 177–186.

Buolamwini J and Gebru T (2018) Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In: *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, *Proceedings of Machine Learning Research*, volume 81. PMLR, pp. 77–91.

Cantal C, Milfont TL, Wilson MS and Gouveia VV (2015) Differential Effects of Right–Wing Authoritarianism and Social Dominance Orientation on Dimensions of Generalized Prejudice in Brazil. *European Journal of Personality* 29(1): 17–27. DOI:10.1002/per.1978.

Carney DR, Jost JT, Gosling SD and Potter J (2008) The secret lives of liberals and conservatives: Personality profiles, interaction styles, and the things they leave behind. *Political psychology* 29(6): 807–840.

Céspedes-Guevara J (2021) A constructionist theory of music induced emotions. DOI:10.31234/osf.io/sfzm2.

Cespedes-Guevara J and Dibben N (2021) Promoting prosociality in Colombia: Is music more effective than other cultural interventions? *Musicae Scientiae* 25(3): 332–357. DOI:10.1177/10298649211013505.

Chen T and Guestrin C (2016) XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)* .

Chen YA, Wang JC, Yang YH and Chen H (2014) Linear Regression-based Adaptation of Music Emotion Recognition Models for Personalization. In: *Proceedings of the IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)*. pp. 2149–2153.

Chen YA, Wang JC, Yang YH and Chen HH (2017) Component Tying for Mixture Model Adaptation in Personalization of Music Emotion Recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 25(7): 1409–1420.

Chomsky N (2004) On Colombia. In: Stokes D (ed.) *Americas's Other War: Terrorizing Colombia*, chapter On Colombia. Zed Books, pp. viii–x.

Crawford K (2021) *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale Univeristy Press.

Céspedes-Guevara J and Eerola T (2018) Music communicates affects, not basic emotions - A constructionist account of attribution of emotional meanings to music. *Frontiers in Psychology* 9(FEB): 1–19.

da Silva Mahleiro RM (2016) *Emotion-based Analysis and Classification of Music Lyrics*. PhD Thesis, Universidade de Coimbra.

Damasio AR (1994) *Descartes' Error: Emotion, Reason, and the Human Brain*. Putnam's Sons.

Delsing MJ, Ter Bogt TF, Engels RC and Meeus WH (2008) Adolescents' music preferences and personality characteristics. *European Journal of Personality: Published for the European Association of Personality Psychology* 22(2): 109–130.

Duckitt J, Bizumic B, Krauss SW and Heled E (2010) A tripartite approach to right-wing authoritarianism: The authoritarianism-conservatism-traditionalism model. *Political Psychology* 31(5): 685–715.

Dunn PG, de Ruyter B and Bouwhuis DG (2012) Toward a better understanding of the relation between music preference, listening behavior, and personality. *Psychology of Music* 40(4): 411–428.

Eder AB and Hommel B (2013) Anticipatory control of approach and avoidance: An ideomotor approach. *Emotion Review* 5(3): 275–279.

Eerola T (2018) Music and Emotion. In: Bader R and Koelsch S (eds.) *Handbook of Systematic Musicology*, chapter Music and Emotion. Springer, pp. 539–556.

Ellsworth PC and Scherer KR (2003) *Appraisal processes in emotion*, chapter Appraisal processes in emotion. Oxford University Press Oxford, pp. 572–595.

European Commission E, for Justice DG, Consumers, Lupiáñez-Villanueva F, Boluda A, Bogliacino F, Liva G, Lechardoy L and Rodríguez de las Heras Ballell T (2022) *Behavioural study on unfair commercial practices in the digital environment : dark patterns and manipulative personalisation*. European Commission. DOI:doi/10.2838/859030.

Eyben F, Weninger F, Gross F and Schuller B (2013) Recent Developments in OpenSMILE, the Munich Open-source Multimedia Feature Extractor. In: *Proceedings of the 21st ACM International Conference on Multimedia*. New York, NY, USA, pp. 835–838.

Fabbri F, Wang Y, Bonchi F, Castillo C and Mathioudakis M (2022) Rewiring What-to-Watch-Next Recommendations to Reduce Radicalization Pathways. In: *Proceedings of the ACM*

Web Conference 2022, WWW '22. New York, NY, USA: Association for Computing Machinery, p. 2719–2728. DOI: 10.1145/3485447.3512143.

Fals Borda O et al. (2001) *Movimientos Sociales, Estado y Democracia en Colombia*. Universidad Nacional de Colombia.

Feldmann AE (2019) *Colombia's Polarizing Peace Efforts*, chapter 6. The Global Challenge of Political Polarization. Brookings Institution Press, pp. 153–176. DOI:10.7864/j.ctvbd8j2p.9.

Fischer P and Greitemeyer T (2006) Music and aggression: The impact of sexual-aggressive song lyrics on aggression-related thoughts, emotions, and behavior toward the same and the opposite sex. *Personality and Social Psychology Bulletin* 32(9): 1165–1176. DOI:10.1177/0146167206288670.

Frijda NH (2007) *The laws of emotion*. Psychology Press.

Frijda NH et al. (1986) *The emotions*. Cambridge University Press.

Gabrielsson A (2006) Emotion perceived and emotion felt: Same and different. *Musicae Scientiae* 10(2): 191–213.

Garrett RK (2009) Echo chambers online?: Politically motivated selective exposure among Internet news users. *Journal of Computer-Mediated Communication* 14(2): 265–285. DOI: 10.1111/j.1083-6101.2009.01440.x.

Gómez-Cañón JS, Herrera P, Cano E and Gómez E (2021) Personalized musically induced emotions of not-so-popular Colombian music. In: *Human-centered Artificial Intelligence Workshop at the 35th Conference on Neural Information Processing Systems (NeurIPS)*. pp. 1–7.

Grajales J (2021) *Agrarian Capitalism, War and peace in Colombia*. Routledge.

Greitemeyer T (2011) Exposure to music with prosocial lyrics reduces aggression: First evidence and test of the underlying mechanism. *Journal of Experimental Social Psychology* 47(1): 28–36. DOI:10.1016/j.jesp.2010.08.005.

Gómez-Cañón JS, Cano E, Eerola T, Herrera P, Hu X, Yang YH and Gómez E (2021a) Music Emotion Recognition: towards new robust standards in personalized and context-sensitive applications. *IEEE Signal Processing Magazine* 38.

Gómez-Cañón JS, Cano E, Yang YH, Herrera P and Gómez E (2021b) Let's agree to disagree: Consensus Entropy Active Learning for Personalized Music Emotion Recognition. In: *Proceedings of the 22nd International Society for Music Information Retrieval Conference (ISMIR)*. pp. 237–245.

Gómez-Cañón JS, Gutiérrez-Páez N, Porcaro L, Porter A, Cano E, Herrera P, Gkiokas A, Santos P, Hernández-Leo D, Karreman C and Gómez E (2022) TROMPA-MER: an open data set for personalized Music Emotion Recognition. *IEEE Transactions on Affective Computing (submitted)* .

Hansen F and Christensen S (2007) *Emotions, Advertising and Consumer Choice*. Copenhagen Business School Press.

Heisbourg E and Feitosa F (2021) Does music affect citizens' evaluations of candidates? *Politics and Life Sciences* 40(2): 172–178. DOI:10.1017/pls.2021.22.

Herrera P (2009) Música y persuasión. In: Ayats J et al. (eds.) *La música y su reflejo en la sociedad*, chapter Música y persuasión. Indigestió Musical, pp. 27–38.

Ho AK, Sidanius J, Kteily N, Sheehy-Skeffington J, Pratto F, Henkel KE, Foels R and Stewart AL (2015) The nature of social dominance orientation: Theorizing and measuring preferences for intergroup inequality using the new SDO7 scale. *Journal of Personality and Social Psychology* 109(6): 1003.

Hoerger M (2010) Participant dropout as a function of survey length in internet-mediated university studies: Implications for study design and voluntary participation in psychological research. *Cyberpsychology, behavior, and social networking* 13(6): 697–700.

Hsueh PY, Melville P and Sindhwani V (2009) Data Quality from Crowdsourcing: A Study of Annotation Selection Criteria. In: *Proceedings of the NAACL HLT 2009 Workshop on Active Learning for Natural Language Processing*, HLT '09. USA: Association for Computational Linguistics, p. 27–35.

Hu X, Chen J and Wang Y (2021) University students' use of music for learning and well-being: A qualitative study and design implications. *Information Processing and Management* 58(1): 1–14.

Hu X, Li F and Ng J (2018) On the Relationships between Music-induced Emotion and Physiological Signals. In: *Proceedings of the 19th International Society for Music Information Retrieval Conference*. Paris, France, pp. 362–369.

Huang R, Sturm BLT and Holzapfel A (2021) De-centering the West: East Asian Philosophies and the Ethics of Applying Artificial Intelligence to Music. In: *Proceedings of the 22nd International Society for Music Information Retrieval Conference (ISMIR)*. Online, pp. 301–309. DOI:10.5281/zenodo.5624543.

Huron D (2016) *Music and meaning*, chapter 15. 2 edition. Oxford University Press Oxford, pp. 233–245.

Jurafsky D and Martin JH (2009) *Speech and Language Processing*. Prentice-Hall, Inc.

Juslin PN (2013) From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews* 10: 235–266. DOI:10.1016/j.plrev.2013.05.008.

Juslin PN (2019) *Musical Emotions Explained*. Oxford: Oxford University Press.

Juslin PN, Barradas G and Eerola T (2015) From Sound to Significance: Exploring the Mechanisms Underlying Emotional Reactions to Music. *The American Journal of Psychology* 128(3): 281–304.

Juslin PN and Västfjäll D (2008) Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and brain sciences* 31(5): 559–575.

Kant I (1790) *Critique of the Power of Judgment*. Cambridge University Press.

Katz-Rosene J (2017) *From Protest Song to Social Song: Music and Politics in Colombia 1966-2016*. PhD Thesis, City University of New York.

Kivy P (1990) *Music alone: reflections on a purely musical experience*. Cornell University Press.

Kuppens P, Champagne D and Tuerlinckx F (2012) The dynamic interplay between appraisal and core affect in daily life. *Frontiers in psychology* 3: 380–388.

Laurier C (2011) *Automatic Classification of Musical Mood by Content-Based Analysis*. PhD Thesis, Universitat Pompeu Fabra.

Laurier C and Herrera P (2009) Automatic detection of emotion in music: Interaction with emotionally sensitive machines. In: Vallverdú J and Casacuberta D (eds.) *Handbook of Research on Synthetic Emotions and Sociable Robotics: New Applications in Affective Computing and Artificial Intelligence*, chapter

Automatic detection of emotion in music: Interaction with emotionally sensitive machines. IGI Global, pp. 9–33.

Lennie TM and Eerola T (2022) The CODA Model: A Review and Skeptical Extension of the Constructionist Model of Emotional Episodes Induced by Music. *Frontiers in Psychology* 13. DOI: 10.3389/fpsyg.2022.822264.

Mahoney RD (2020) *Colombia: what everybody needs to know*. Oxford University Press.

Manning CD and Schütze H (1999) *Foundations of Statistical Natural Language Processing*. Cambridge, MA: MIT Press.

Mcpherson M, Smith-Lovin L and Cook JM (2001) BIRDS OF A FEATHER: Homophily in Social Networks. *Annual Review of Sociology* 27: 415–444.

Mehr SA, Schachner A, Katz RC and Spelke ES (2013) Two Randomized Trials Provide No Consistent Evidence for Nonmusical Cognitive Benefits of Brief Preschool Music Enrichment. *PLOS ONE* 8(12): 1–12.

Mohamed S, Png MT and Isaac W (2020) Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence. *Philosophy & Technology* 33: 659–684. DOI: 10.1007/s13347-020-00405-8.

Moors A (2017) Integration of Two Skeptical Emotion Theories: Dimensional Appraisal Theory and Russell's Psychological Construction Theory. *Psychological Inquiry* 28(1): 1–19. DOI: 10.1080/1047840X.2017.1235900.

Moors A, Boddez Y and De Houwer J (2017) The power of goal-directed processes in the causation of emotional and other actions. *Emotion Review* 9(4): 310–318.

Moors A, Ellsworth PC, Scherer KR and Frijda NH (2013) Appraisal theories of emotion: State of the art and future development. *Emotion Review* 5(2): 119–124.

Moors A and Kuppens P (2008) Distinguishing between two types of musical emotions and reconsidering the role of appraisal. *Behavioral and Brain Sciences* 31(5): 588–589.

Nigam K, Mccallum AK, Thrun S and Mitchell T (2000) Text Classification from Labeled and Unlabeled Documents using EM. *Machine Learning* 39(2): 103–134. DOI:10.1023/A:1007692713085.

Noble SU (2018) *Algorithms of oppression: How search engines reinforce racism*. New York University Press.

Ollen JE (2006) *A criterion-related validity test of selected indicators of musical sophistication using expert ratings*. PhD Thesis, Ohio State University.

O'Neil C (2016) *Weapons of Math Destruction*. Crown.

Orozco Perez HD, Dumas G and Lehmann A (2020) Binaural Beats through the Auditory Pathway: From Brainstem to Connectivity Patterns. *eNeuro* 7(2).

Palazzi A, Wagner Fritzen B and Gauer G (2019) Music-induced emotion effects on decision-making. *Psychology of Music* 47(5): 621–643. DOI:10.1177/0305735618779224.

Panda R, Malheiro RM and Paiva RP (2020) Audio Features for Music Emotion Recognition: a Survey. *IEEE Transactions on Affective Computing* Accepted for publication: 1–20. DOI: 10.1109/TAFFC.2020.3032373.

Panda R, Rui RM and Paiva P (2018) Musical texture and expressivity features for music emotion recognition. In: *Proceedings of the 19th International Society for Music Information Retrieval Conference*. Paris, France, pp. 383–391.

Pariser E (2011) *The filter bubble: What the internet is hiding from you*. Penguin UK.

Pratto F, Sidanius J, Stallworth LM and Malle BF (1994) Social Dominance Orientation: A Personality Variable Predicting Social and Political Attitudes. *Journal of Personality and Social Psychology* 67(4): 741–763. DOI:10.1037/0022-3514.67.4.741.

Quishpe R (2020) Corcheas insurgentes: usos y funciones de la música de las FARC-EP durante el conflicto armado en Colombia. *Izquierdas* 49: 554–579. DOI:10.4067/s0718-50492020000100231.

Rentfrow PJ (2012) The role of music in everyday life: Current directions in the social psychology of music. *Social and personality psychology compass* 6(5): 402–416.

Rentfrow PJ, Goldberg LR and Zilca R (2011) Listening, watching, and reading: The structure and correlates of entertainment preferences. *Journal of personality* 79(2): 223–258.

Rentfrow PJ and Gosling SD (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.

Rosenberg RE (2021) Perfect Pitch: 432 Hz Music and the Promise of Frequency. *Journal of Popular Music Studies* 33(1): 137–154. DOI:10.1525/jpms.2021.33.1.137.

Russell JA (1980) A Circumplex Model of Affect. *Personality and Social Psychology* 39(6): 1161–1178.

Russell JA (2003) Core affect and the psychological construction of emotion. *Psychological review* 110(1): 145–172.

Ruth N (2018) "They don't really care…": Effects of music with prosocial content and corresponding media coverage on prosocial behavior. *Musicae Scientiae* 22(3): 415–433.

Ruth N and Schramm H (2021) Effects of prosocial lyrics and musical production elements on emotions, thoughts and behavior. *Psychology of Music* 49(4): 759–776. DOI:10.1177/0305735620902534.

Saarikallio S (2019) *Music as a resource for agency and empowerment in identity construction*, chapter 8. Oxford University Press Oxford, pp. 89–98.

Saarikallio S, Gold C and McFerran K (2015) Development and validation of the Healthy-Unhealthy Music Scale. *Child and Adolescent Mental Health* 20(4): 210–217. DOI:https://doi.org/10.1111/camh.12109.

Sarasúa Á, Laurier C and Herrera P (2012) Support Vector Machine Active Learning for Music Mood Tagging. In: *Proceedings of the 9th International Symposium on Computer Music Modelling and Retrieval (CMMR)*. London, England, pp. 518–525.

Saunders BA and Ngo J (2017) *The Right-Wing Authoritarianism Scale*, chapter The Right-Wing Authoritarianism Scale. Cham: Springer International Publishing, pp. 1–4. DOI:10.1007/978-3-319-28099-8_1262-1.

Schedl M, Flexer A and Urbano J (2013) The neglected user in music information retrieval research. *Journal of Intelligent Information Systems* 41: 523–539.

Scherer K and Zentner M (2008) Music evoked emotions are different–more often aesthetic than utilitarian. *Behavioral and Brain Sciences* 31(5): 595–596.

Scherer KR (1982) Emotion as a process: Function, origin and regulation. *Social Science Information* 21(4-5): 555–570. DOI: 10.1177/053901882021004004.

Scherer KR (2009a) Emotions are emergent processes: they require a dynamic computational architecture. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364(1535): 3459–3474.

Scherer KR (2009b) The dynamic architecture of emotion: Evidence for the component process model. *Cognition and emotion* 23(7): 1307–1351.

Scherer KR and Coutinho E (2013) How music creates emotion: a multifactorial process approach. *The emotional power of music: Multidisciplinary perspectives on musical arousal, expression, and social control* : 121–145.

Scherer KR and Moors A (2019) The emotion process: Event appraisal and component differentiation. *Annual review of psychology* 70(1): 719–745.

Schiller D (2022) The human affectome. DOI:10.31234/osf.io/9nu32. URL psyarxiv.com/9nu32.

Schuller BW (2013) *Intelligent audio analysis*. Springer.

Schölkopf B, Platt J and Hofmann T (2007) Correcting Sample Selection Bias by Unlabeled Data. In: *Advances in Neural Information Processing Systems 19: Proceedings of the 2006 Conference*. The MIT Press, pp. 601–608. DOI:10.7551/mitpress/7503.003.0080.

Sharman L and Dingle GA (2015) Extreme Metal Music and Anger Processing. *Frontiers in Human Neuroscience* 9. DOI:10.3389/fnhum.2015.00272.

Shepherd D and Sigg N (2015) Music preference, social identity, and self-esteem. *Music Perception: An Interdisciplinary Journal* 32(5): 507–514.

Silverman MJ, Gooding LF and Yinger O (2020) It's...complicated: A theoretical model of music-induced harm. *Journal of Music Therapy* 57(3): 251–281. DOI:10.1093/jmt/thaa008.

Sloboda JA and Juslin PN (2010) *At the Interface Between the Inner and Outer World*, chapter 4. Oxford University Press, pp. 73–99.

Sloman SA (1996) The empirical case for two systems of reasoning. *Psychological bulletin* 119(1): 3–22.

Smith CA and Lazarus RS (1991) *Emotion and adaptation*, chapter ? Guilford Press, pp. 609–637.

Stark L (2016) The emotional context of information privacy. *The Information Society* 32(1): 14–27. DOI:10.1080/01972243.2015.1107167.

Steele JR and Brown JD (1995) Adolescent room culture: Studying media in the context of everyday life. *Journal of youth and adolescence* 24(5): 551–576.

Stokes D (2005) *America's other war: Terrorizing Colombia*. Zed Books.

Stray J (2021) Designing Recommender Systems to Depolarize. *First Monday* .

Su D and Fung P (2012) Personalized Music Emotion Classification via Active Learning. In: *Proceedings of the Second International ACM Workshop on Music Information Retrieval with User-Centered and Multimodal Strategies*. New York, NY, USA, p. 57–62.

Sunstein CR (2002) The Law of Group Polarization. *Journal of Political Philosophy* 10(2): 175–195. DOI:10.1111/1467-9760.00148.

Tarrant M (2002) Adolescent peer groups and social identity. *Social Development* 11(1): 110–123. DOI:10.1111/1467-9507.00189.

Tarrant M, North AC and Hargreaves DJ (2002) Youth identity and music.

Ter Bogt TF, Engels RC, Bogers S and Kloosterman M (2010) "Shake it baby, shake it": Media preferences, sexual attitudes and gender stereotypes among adolescents. *Sex roles* 63(11): 844–859.

Thompson WF, Quinto L et al. (2011) Music and emotion: Psychological considerations. *The aesthetic mind: Philosophy and psychology* : 357–375.

Tucker JA, Guess A, Barberá P, Vaccari C, Siegel A, Sanovich S, Stukal D and Nyhan B (2018) Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature. *SSRN* .

Valero FB, Baranes M and Epure EV (2022) Topic Modeling on Podcast Short-Text Metadata. In: *44th European Conference on Information Retrieval (ECIR)*. pp. 1–16.

Véliz C (2020) *Privacy Is Power: Why and How You Should Take Back Control of Your Data*. Bantam Press.

Villalobos J and Ramírez-Pimienta J (2004) Corridos and la pura verdad: Myths and Realities of the Mexican Ballad. *South Central Review* 21: 129–149. DOI:10.1353/scr.2004.0050.

Warrenburg LA (2020) Comparing musical and psychological emotion theories. *Psychomusicology: Music, Mind, and Brain* 30(1): 1–19.

Weninger F, Eyben F, Schuller BW, Mortillaro M, Scherer KR and Krajewski J (2013) On the acoustics of emotion in audio: what speech, music, and sound have in common. *Frontiers in Psychology* 4: 1–12. DOI:10.3389/fpsyg.2013.00292.

Wu X, Li C, Zhu Y and Miao Y (2020) Short Text Topic Modeling with Topic Distribution Quantization and Negative Sampling Decoder. In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Online: Association for Computational Linguistics, pp. 1772–1782. DOI:10.18653/v1/2020.emnlp-main.138.

Yan X, Guo J, Lan Y and Cheng X (2013) A Biterm Topic Model for Short Texts. In: *Proceedings of the 22nd International Conference on World Wide Web*, WWW '13. New York, USA: Association for Computing Machinery, p. 1445–1456. DOI:10.1145/2488388.2488514.

Yang YH and Chen HH (2011) *Music Emotion Recognition*. CRC Press.

Yang YH, Su YF, Lin YC and Chen HH (2007) Music Emotion Recognition: The Role of Individuality. In: *Proceedings of the International Workshop on Human-centered Multimedia*. pp. 13–22.

Yin J and Wang J (2014) A Dirichlet Multinomial Mixture Model-Based Approach for Short Text Clustering. In: *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '14. New York, USA: Association for Computing Machinery, p. 233–242. DOI:10.1145/2623330.2623715.

Zakrisson I (2005) Construction of a short version of the Right-Wing Authoritarianism (RWA) scale. *Personality and individual differences* 39(5): 863–872.

Zamosc L (1986) *The agrarian question and the peasant movement in Colombia*. Cambridge University Press.

Zangerle E, Chen CM, Tsai MF and Yang YH (2021) Leveraging Affective Hashtags for Ranking Music Recommendations. *IEEE Transactions on Affective Computing* 12(1): 78–91. DOI:10.1007/978-981-15-7511-2_76.

Zhang JD and Schubert E (2019) A Single Item Measure for Identifying Musician and Nonmusician Categories Based on Measures of Musical Sophistication . *Music Perception* 36(5): 457–467. DOI:10.1525/MP.2019.36.5.457.

Ziv N (2016) Music and compliance: Can good music make us do bad things? *Psychology of Music* 44(5): 953–966. DOI: 10.1177/0305735615598855.

Zuboff S (2019) *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Public Affairs.

Zweigenhaft RL (2008) A do re mi encore: A closer look at the personality correlates of music preferences. *Journal of individual differences* 29(1): 45–55.