

This is a repository copy of *Popular Music as Entertainment Communication: How Perceived Semantic Expression Explains Liking Previously Unknown Music*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/163797/>

Version: Published Version

---

**Article:**

Lepa, Steffen, Steffens, Jochen, Herzog, Martin et al. (1 more author) (2020) Popular Music as Entertainment Communication: How Perceived Semantic Expression Explains Liking Previously Unknown Music. Media and Communication. ISSN: 2183-2439

<https://doi.org/10.17645/mac.v8i3.3153>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

Article

# Popular Music as Entertainment Communication: How Perceived Semantic Expression Explains Liking of Previously Unknown Music

Steffen Lepa <sup>1,\*</sup>, Jochen Steffens <sup>2</sup>, Martin Herzog <sup>1</sup> and Hauke Egermann <sup>3</sup>

<sup>1</sup> Audio Communication Group, TU Berlin, 10587 Berlin, Germany; E-Mail: steffen.lepa@tu-berlin.de (S.L.), herzog@tu-berlin.de (M.H.)

<sup>2</sup> Media Department, University of Applied Sciences Düsseldorf, 40476 Düsseldorf, Germany; E-Mail: jochen.steffens@hs-duesseldorf.de

<sup>3</sup> York Music Psychology Group, University of York, York, YO10 5DD, UK; E-Mail: hauke.egermann@york.ac.uk

\* Corresponding author

Submitted: 14 April 2020 | Accepted: 26 June 2020 | Published: 13 August 2020

## Abstract

Our contribution addresses popular music as essential part of media entertainment offerings. Prior works explained liking for specific music titles in ‘push scenarios’ (radio programs, music recommendation, curated playlists) by either drawing on personal genre preferences, or on findings about ‘cognitive side effects’ leading to a preference drift towards familiar and society-wide popular tracks. However, both approaches do not satisfactorily explain why previously unknown music is liked. To address this, we hypothesise that unknown music is liked the more it is perceived as emotionally and semantically expressive, a notion based on concepts from media entertainment research and popular music studies. By a secondary analysis of existing data from an EU-funded R&D project, we demonstrate that this approach is more successful in predicting 10000 listeners’ liking ratings regarding 549 tracks from different genres than all hitherto theories combined. We further show that major expression dimensions are perceived relatively homogeneous across different sociodemographic groups and countries. Finally, we exhibit that music is such a stable, non-verbal sign-carrier that a machine learning model drawing on automatic audio signal analysis is successfully able to predict significant proportions of variance in musical meaning decoding.

## Keywords

entertainment; genre preferences; musical expression; music preferences; musical taste; popular music; push scenarios; semantics

## Issue

This article is part of the issue “Computational Approaches to Media Entertainment Research” edited by Johannes Breuer (GESIS—Leibniz Institute for the Social Sciences, Germany), Tim Wulf (LMU Munich, Germany) and M. Rohangis Mohseni (TU Ilmenau, Germany).

© 2020 by the authors; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

## 1. Introduction

Popular music (in the broadest sense, also encompassing ‘oldies,’ jazz and hits from the classical repertoire, based on the definition of Tagg, 2000), is one of the most prevalent types of entertainment content in everyday media use, especially in social media. It is nowadays predominantly consumed in *push scenarios*—socio-musical con-

texts, in which music is selected and played back for us by someone else (e.g., when listening to radio programs, curated playlists, DJ sets, in-store music, YouTube videos, shuffle-mode, and music in virtual worlds) or by recommendation algorithms. The current abundance of listening situations where people are confronted with previously unknown popular music is part of the ongoing “musicalization” of society (Pontara & Volgsten, 2017).

In the past age of audio storage media, the breadth of existing music tracks available for airplay was generally limited by material physical and economic restrictions, similar to the number and stylistic variety of albums and singles published and sold. Consequently, the question of which records to buy and which musical genre and artists to adhere to has led to a great deal of socio-cultural distinction practices (Bourdieu, 1984). Out of limited and heterogeneous economic and cultural capital, people typically stuck to their cultural habitus acquired during their formative years which then formed an essential part of their identity during their later lives (Frith, 1996). Therefore, personal taste and choices in music selection have always tended to separate people of different generations, milieus, and cultures from one another, which remains partly the case today (Mellander, Florida, Rentfrow, & Potter, 2018; Vlegels & Lievens, 2017). However, we claim that due to the inflation of musico-technological repertoires in the age of digital media (Lepa & Hoklas, 2015) as well as the global introduction of 'flatrate' streaming offerings (Drott, 2018) and music recommendation algorithms (Krämer, 2018), there is a growing tendency that the logics governing people's socio-musical practices are converging to new patterns (see a review in Section 1.1). In consequence, theoretical models from music psychology and cultural sociology that successfully described musical preference dynamics in the past require re-examination (Brisson & Bianchi, 2019; Prior, 2013).

Accordingly, in the present article, we propose and empirically compare alternative explanations for music liking in existing popular music push scenarios by drawing on concepts from empirical aesthetics, media entertainment research, and popular music studies.

### 1.1. Challenging the Perceived View on Music Liking: Personal Genre Preferences

Research on music liking has often employed questionnaires asking participants for the degree of liking regarding several given musical genre labels. This practice has revealed sociodemographic differences in obtained preference patterns (typically, a high-brow vs low-brow cultural gap; Roose & Stichele, 2010), as well as correlations with personality traits (Fricke & Herzberg, 2017) and political attitudes (Feezell, 2017). However, observed effect sizes are comparatively small (Schäfer & Mehlhorn, 2017), and it remains unclear whether obtained answer patterns relate to music listening practices in the push scenarios under discussion here. This is due to the ambiguous intensional content of musical genres. In catalogues of music stores or streaming providers, different taxonomies exist (e.g., Spotify vs Apple Music), and listeners tend to have heterogeneous and historically changing ideas about the musical attributes and values defining specific genres (Lahire, 2008). Moreover, non-musicians in particular associate genres rather with social stereotypes and identity concepts related to artists, epochs, subcultures, and fandom the music stems from

(Shevy, 2008). As a result, genre-based expressed musical taste has become an elemental part of postmodern identity and distinction practices (Lonsdale & North, 2009).

However, while being an interesting phenomenon in itself, taste performances (Hennion, 2001) are not necessarily informative for the actual patterns of music liking and listening practice (Lonsdale & North, 2012). Also, recent empirical studies suggest a growing tendency towards genre "omnivorousness" (Peterson, 1992) spreading across social classes (Vlegels & Lievens, 2017) and a continuous development towards new genre taxonomy logics based on social context or lifeworld functions (Airoldi, Beraldo, & Gandini, 2016). It is therefore unsurprising that the actual power of traditional genre labels for predicting musical liking is rather low (Brisson & Bianchi, 2019). Hence, we infer that explicitly stated genre affinities might only explain a minor portion of music liking in music push scenarios.

### 1.2. Familiarity, Prominence and Popularity as 'Cognitive Side Effects' in Music Liking

In contrast, theories from empirical aesthetics and social psychology appear better suited to explaining music liking in times of musicalization. For instance, repeated exposure to a stimulus leads to cognitive fluency effects (Jacoby & Dallas, 1981) and a more positive evaluation. However, familiarity and pleasantness of artworks only covary up to a certain ideal point, from where pleasantness decreases again in terms of a saturation effect, often idealised graphically by an inverted U-curve (Chmiel & Schubert, 2017). In the original theory of Berlyne (1971), this effect was said to interact with stimulus complexity. However, this has rarely been successfully demonstrated in experimental works with musical stimuli (Madison & Schiölde, 2017). An alternative explanation of the advantage of widely-known music in push scenarios is the elaboration likelihood model (Petty & Cacioppo, 1986). According to the model, the fact that a piece of music is well-known and appreciated by others (e.g., in our culture or peer-group) constitutes a peripheral persuasive cue that can positively influence aesthetic experiences, in particular in the low-involvement scenarios that we discuss here (Egermann, Grewe, Kopiez, & Altenmüller, 2009). Overall, we assume personal *familiarity* and social *popularity* effects in combination with incidental *saturation* resulting from over-prominence can explain experienced liking in situations where we are confronted with familiar-sounding music. We will denote them throughout this paper as "cognitive side effects" because they affect music liking independently of actual musical content or perceived expression.

### 1.3. Musical Expression Strength and Breadth as New Explanation for Music Liking

Rentfrow, Goldberg, and Levitin (2011) and Rentfrow et al. (2012), criticizing the genre label-based approach,

suggested working with sounding questionnaires to operationalize musical preferences. Following Hevner (1936), they also introduced adjective inventories allowing listeners to describe perceived ‘attributes of music.’ Based on aesthetical judgements gathered in this way, Greenberg et al. (2016) identified three major dimensions of perceived musical expression, two of them representing affect (“Arousal” & “Valence”) and one representing the felt degree of aesthetic-cognitive stimulation (“Depth”). While this has generated significant progress for the field of music liking research, choosing a rather small convenience sample for ‘judging’ the semantics of popular music may lead to a narrowing of possible meanings as to what is deemed valuable from a high-brow perspective. This might lead to the perspective of the ‘people,’ the actual producers and addressees of popular music (Frith, 1996), becoming neglected.

Furthermore, analogous to discussions in media entertainment research (Klimmt, 2011), it appears crucial to acknowledge that beyond affect-guided hedonism and intellectual appreciation, people might also enjoy specific types of music because they fulfil their eudaimonic needs and help them to find identity, truth, and transformational experiences, overall rendering their everyday existence meaningful (Vorderer & Reinecke, 2015). To explain how meaning is imparted, Tagg (2013) argues that a majority of meanings conveyed by popular music, including substantial parts of affect expression, are due to so-called “para-musical fields of connotation.” This term describes extra-musical meanings that are bestowed upon musical sign-carriers by human appropriation practices during the music’s semiotic carrier as part of the circulation of culture (Herzog, Lepa, Egermann, Schönrock, & Steffens, 2020).

Based on the theoretical perspective of non-verbal communication theory adopted for music (Brunswik, 1952; Juslin, 2000) we further assume that perceived musical expression is only partly idiosyncratic, and rather by and large pragmatically ‘understood’ homogenously by other recipients, because most of our conspecifics take part in the same cultural game of musical semiosis as we do. In empirical studies on music expression drawing on Brunswik’s (1952) lens model (Eerola, Friberg, & Bresin, 2013; Juslin, 2000), it has been found that the communicative cues employed in music work in a linear-additive fashion, sometimes redundantly, but often imparting several dimensions of meaning at the same time, which renders some pieces so expressive and popular. Hence, we postulate that liking for previously unknown popular music is dependent on the *breadth and strength* of perceived musical expression, which, according to cultural theorist Alison Stone (2016) should encompass the dimensions of affect, values, aesthetic commitments, identity, location and time. We furthermore suspect that these decoded connotations do exhibit a certain degree of idiosyncratic and cultural heterogeneity (Kristen & Shevy, 2013) in terms of content and their specific weight in personal preference judgments.

#### 1.4. The Constructivist Challenge: To Which Degree Are Perceived Musical Expressions Socially Uniform and Predictable?

The current lack of prior systematic research on semantic musical expression might stem from the ‘constructivist challenge’ imposed by the concept of para-musical fields of connotations. If musical meaning lies to a greater extent in the ‘ear of the beholder’ and is not immediately inherent to the acoustic stimulus, how can we systematically measure it? On the other hand, it is known that film scores and advertisement music work well in communicating certain connotations successfully to recipients (Bouvier & Machin, 2013). Furthermore, it should be considered that musicalization has probably already led to a perceptible degree of musical sign-disambiguation across the globe and emotional music expressions might be based to some degree on cross-cultural universals (Sievers, Polansky, Casey, & Wheatley, 2012). Finally, similar to story or movie interpretations, the empirically found extent of non-uniformity in meaning decoding might also be due to the specificity of meanings searched for (Lepa, 2010). Following Tagg (2013), the actual degree of non-uniformity could be analysed either by measuring the variance of a small audience’s actual meaning productions regarding a smaller pool of music or by formalising human meaning attribution regarding a larger pool of music with machine learning (ML) methods and then checking the resulting explanatory model power when applied to new music. Both approaches are pursued in the current contribution.

#### 1.5. Resulting Hypotheses and Research Questions

Due to participation in a multi-national European research and development project on music branding funded by the EU ([www.abc\\_dj.eu](http://www.abc_dj.eu)), we had the opportunity to test some of the assumptions mentioned above with an existing dataset. Even if the actual expression potential of popular music most probably reaches beyond the commercially exploitable domain, this nevertheless provided an excellent opportunity to test our following theoretical hypotheses based on the possibilities of this specific dataset:

H1: Liking of presented music is (positively) dependent on personal *genre affinity* strength.

H2: Liking of presented music is (positively) dependent on personal *familiarity* with a track.

H3: Liking of presented music is dependent on society-wide *popularity* and dependent on society-wide *prominence* of a track.

H4: Liking of presented music is dependent on *strength* and *breadth* of perceived musical expression regarding affect and values.

Additionally, four overarching open-ended research questions were explicated that address the constructivist challenge of music semantics:

RQ1: What is the relative importance of hypothesized predictors (and their sub-dimensions) regarding liking of presented music?

RQ2: Are there socio-cultural differences in perceived musical expression or relative preferences for different dimensions of musical expression?

RQ3: To what extent is it possible to predict perceived musical expression based on algorithmic audio signal analysis?

RQ4: Which acoustical attributes of popular music are best suited to predict perceived musical expression dimensions?

Two empirical studies were conducted. Based on a secondary analysis of existing data, Study 1 addresses the four main hypotheses, as well as RQ1 and RQ2. In order to answer RQ3 and RQ4, Study 2 then employs numerical results from Study 1 and combines them with ML and music information retrieval (MIR) techniques.

## 2. Study 1: Explaining Music Liking in Push Scenarios

To perform systematic inquiry on hypotheses H1–H4, as well as RQ1 and RQ2, we drew on available data from a cross-national online survey experiment which was part of an EU Horizon 2020 research & development project (Herzog, Lepa, & Egermann, 2016; Herzog, Lepa, Steffens, Schönrock, & Egermann, 2017a).

### 2.1. Methods

Due to space limitations, details on participants, sampling, stimulus material and data pre-processing are documented in the Supplementary File (A1.1.–A1.3). The resulting net sample comprises  $n = 9,197$  subjects from three generations (gen Y: age 18–34, gen X: age 35–51, gen B: age 52–68) and three countries, with gender being approximately equal-distributed.

#### 2.1.1. Procedure

Based on initial sociodemographic screening procedures organised by panel providers, subjects received an online questionnaire formulated in their country's primary language (English, German, Spanish). They conducted a short sound test and were then presented with either four (wave 1) or six (wave 2) randomly selected 30s popular music excerpts from a larger pool (see Section 2.4 for details). Afterwards, they rated the subjectively perceived fit between the music and 15 adjective attributes (GMBI\_15 inventory, see Figure 3) employing a 6-point

scale, as well as the degree of familiarity with and liking for the excerpt. Finally, subjects stated the extent of their general personal affinity to each of 10 different musical genres in the pool, which were presented to them as linguistic labels (see Figure 1).

### 2.2. Results

#### 2.2.1. Personal Genre Affinities versus Actual Liking

Figure 1 provides an overview of obtained genre preference ratings. Highest personal affinities were found for *Pop & Charts*, and *Rock & Punk*, while *Hip Hop & Trap*, *Country & Folk* and *World Music* received the lowest sympathies. Also, we computed a multivariate generalized linear model (cumulative logit link,  $n = 9,197$ ) to test for socio-cultural differences in genre preference patterns revealing various highly significant differences in line with the literature (not documented here), altogether explaining 15%  $R^2$  (Nagelkerke).

We then calculated ordinal Kendall-Tau correlations between stated affinity for a genre and the actual liking of excerpts from that genre in the prior listening experiment, resulting in an average correlation of  $\tau = 0.22$ . Hence, stated affinities appear to be rather weak predictors for actual liking. Furthermore, as depicted by Figure 6 in the Supplementary File, we observed substantial differences in correlation size across genres and country of residence, hinting at cultural heterogeneities in genre label understanding.

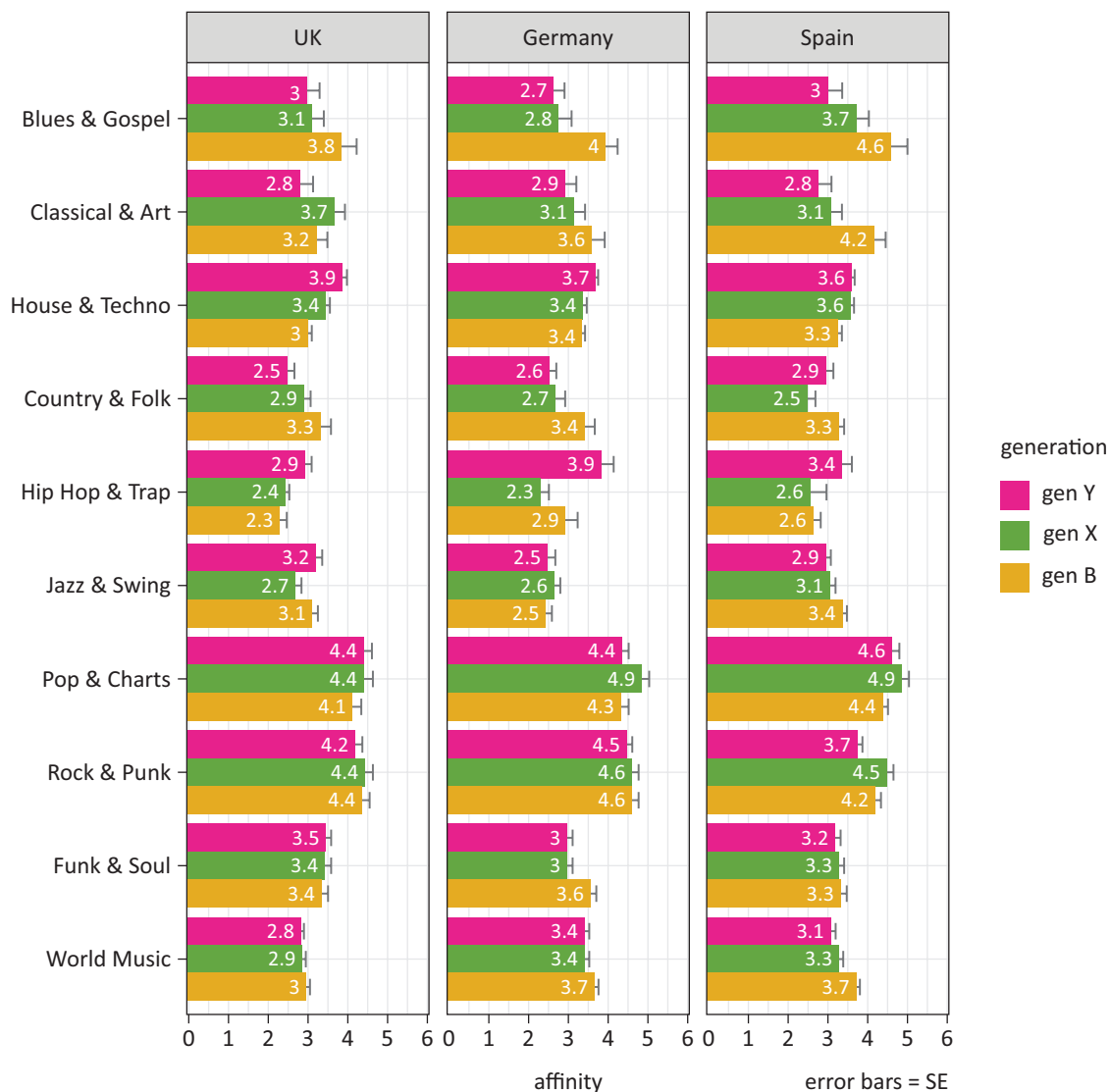
#### 2.2.2. Track Familiarity, Prominence, and Popularity

To test whether there was a sufficient amount of 'novel' music in the pool presented to participants, we inspected histograms of excerpt familiarity ratings by musical genre. Results demonstrated an expected long-tail distribution with the far more frequent *House & Techno* and *World Music* excerpts being rather unfamiliar to respondents, while *Rock & Punk* and *Pop & Charts* were most familiar to them (see Figure 7 in the Supplementary File).

We then calculated a society-wide prominence score for each excerpt, based on the mean track familiarity rating per country. Similarly, we computed a society-wide popularity score for each excerpt, based on the mean track liking rating per country. Afterwards, we estimated the ordinal Kendall-Tau correlation between both measures, resulting in  $\tau = 0.34$ . To check for a possible non-linear dependency, we plotted both aggregated index variables against each other (Figure 2), obtaining a clear 'hinge' effect with weaker dependencies for prominence values above a scale value of 2, but no substantial relationship between both indices and specific musical genres.

#### 2.2.3. Perceived Musical Expression

To measure perceived musical expression, our survey utilised a multi-lingual questionnaire inventory



**Figure 1.** Mean genre affinity by country and generation (scaled from 1 to 6).

(GMBI\_15) that had been developed based on results of an expert focus group and a marketing expert survey (Herzog et al., 2020; Herzog, Lepa, Steffens, Schönrock, & Eggermann, 2017b). The GMBI\_15 measures five musical expression dimensions relevant for branding campaigns, each operationalized by three manifest item indicators. Two dimensions represent musical affect expression (*Arousal*, *Valence*), while three others represent musical value expression (*Authenticity*, *Timeliness*, *Eroticity*). For interpretation of resulting factor scores, it is worth noting that, while items are formulated unipolar, the dimensions of the latent variables are interpreted bipolar (*Arousal*: relaxing–stimulating, *Valence*: dark–bright, *Authenticity*: conventional–authentic, *Timeliness*: traditional–futuristic, *Eroticity*: mental–sensual).

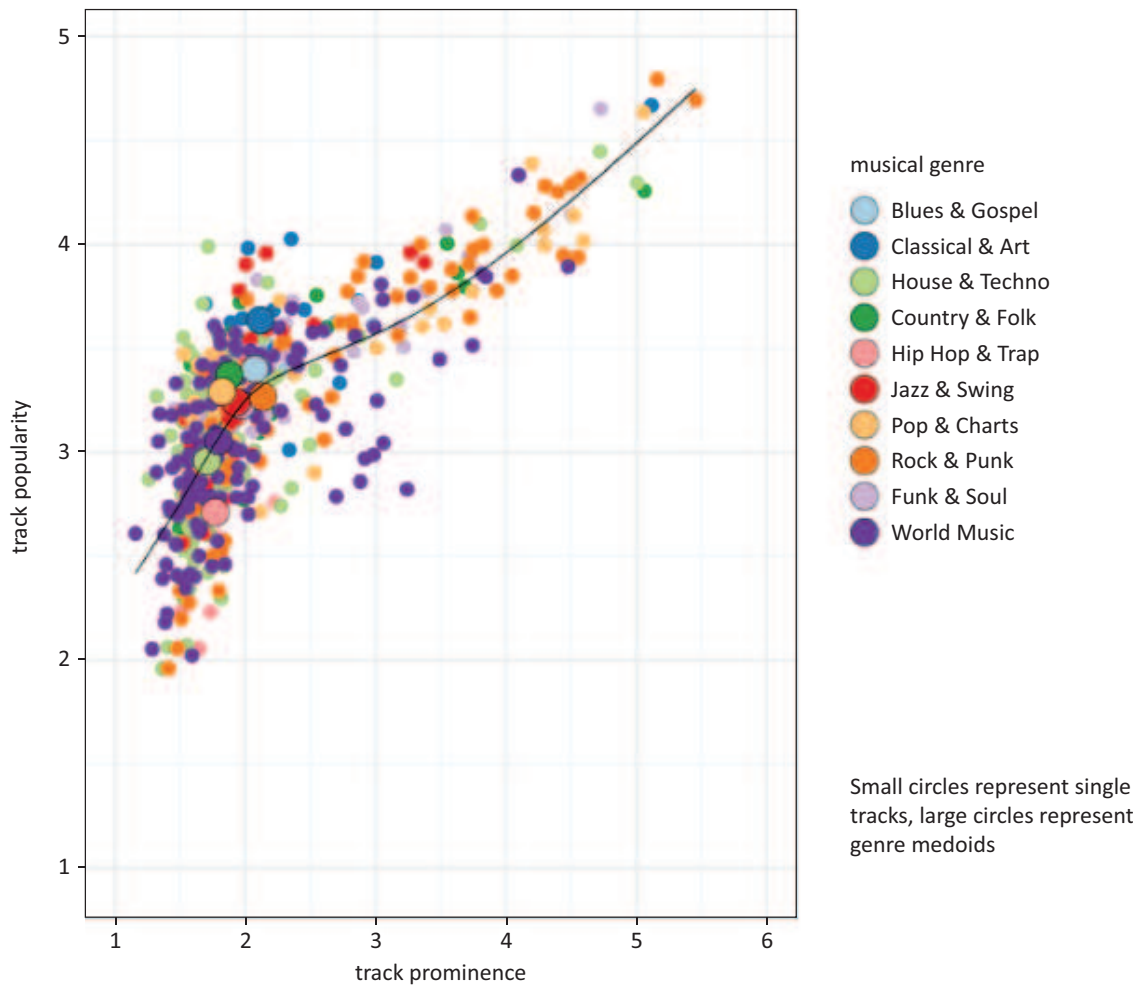
The empirical fit of the employed GMBI\_15 measurement model (Herzog et al., 2017a) was estimated using MLR estimation and a sandwich-estimator to compensate for unbalanced measurement repetition within subjects (see Figure 3). This procedure resulted in a very good fit with  $\chi^2 = 515.239$ ;  $df = 80$ ;

$p < 0.01$ ; RMSEA = 0.040 [0.039–0.041], CFI = 0.968; SRMR = 0.028 (note that significant p-values for model fit are expected for this sample size). Since items had been presented in three different languages, we tested measurement invariance across language versions following Cheung and Rensvold (2002), resulting in a fair degree of scalar invariance (see Table 8 in the Supplementary File). After inverting the polarity of *Arousal* to improve interpretability, we finally calculated z-standardised factor scores for each observation in the dataset. For each musical expression factor, we then performed an ANOVA-based variance component estimation (Searle, 1995), resulting in ~1%  $R^2$  for socio-demographics, while track identity explained between 12–26%  $R^2$  (see Table 9 in the Supplementary File).

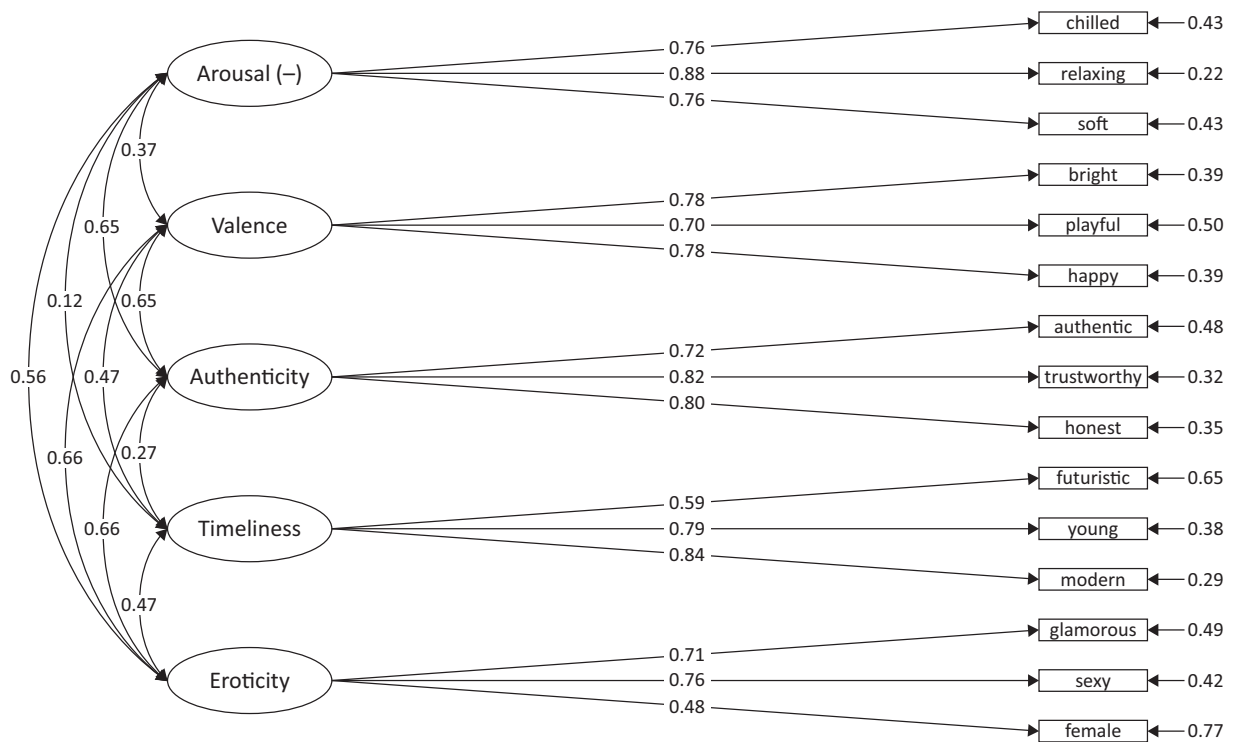
#### 2.2.4. Results of Hypotheses Tests

Hypotheses regarding music liking were tested by a block-wise ordinal logistic regression model (cumulative logit link,  $n = 9,197$ ), calculating cluster-robust standard er-





**Figure 2.** Track prominence vs track popularity by genre.



**Figure 3.** GMBI\_15 measurement model for perceived musical expression (affect/values).

rors to compensate for unbalanced measurement repetitions within subjects. Hypothesis 1 regarding *genre affinity* as well as Hypothesis 3 assuming an influence of *popularity* and *prominence* were deliberately tested as last theoretical model blocks. This was done to allow for estimation of their ‘true’  $R^2$  contribution after having controlled for predictors sharing common variance. The incremental gain in Nagelkerke’s  $R^2$  was calculated for each block corresponding to one of the four major hypotheses (model 1–4), as well as for two additional models (model 5–6), which statistically corrected for the imbalance in track genres and socio-demographics. All hypotheses were confirmed as highly significant, and estimated beta-values for each predictor were only slightly altered when entering controls (Table 1). In an extended model version (not documented here due to space restrictions), we also tested all two-way-interactions between the six expression factor variables and socio-demographics which resulted in some significant, but minor effects contributing to an overall additional  $R^2$  gain of only 1%.

### 2.3. Discussion

Results obtained from Study 1 confirmed our assumption concerning the obsolescence of genre labels for explaining musical liking. In contrast, assumed ‘cognitive side effects’ related to *familiarity*, *popularity*, and *prominence* play an essential role (H2 + H3). Notably, on a societal level, we identified a dampening effect of too much *prominence* on *popularity*. Furthermore, as postulated, it is predominantly music’s perceived expression of affect and values (with nearly similar weights) that explain best why individual people enjoy previously unknown music played back to them (H4). When controlling for these effect clusters, preferences expressed by genre labels only explain a small residual portion of music liking (H1), feasibly representing associated non-musical stereotypes connected with genre labels. Finally, we observed only minor heterogeneities in perceived musical expression across socio-demographics and cultures (RQ1) and, similarly, we only found minor socio-cultural differences in weights for different musical expression dimensions predicting musical liking (RQ2).

**Table 1.** Results of ordinal block-wise regression (cumulative logit link) performed to explain liking of musical excerpts.

Predictor/Model	(1)	(2)	(3)	(4)	(5)	(6)
track familiarity (H2)	0.85***	0.57***	0.54***	0.55***	0.54***	0.54***
expression: arousal (H4)		−0.33***	−0.34***	−0.30***	−0.32***	−0.33***
expression: valence (H4)		0.44***	0.44***	0.47***	0.48***	0.49***
expression: authenticity (H4)		0.70***	0.66***	0.65***	0.62***	0.62***
expression: timeliness (H4)		0.17***	0.13***	0.19***	0.19***	0.18***
expression: eroticity (H4)		0.15***	0.16***	0.08***	0.09***	0.08***
genre affinity (H1)			0.30***	0.31***	0.33***	0.32***
track popularity (H3)				0.51***	0.51***	0.49***
track prominence (H3)				−0.28***	−0.28***	−0.26***
track genre: Blues and Gospel					0.24***	0.25***
track genre: Classical and Art					0.17***	0.17***
track genre: House and Techno					−0.05*	−0.05*
track genre: Country and Folk					0.01	−0.01
track genre: Hip Hop and Trap					0.21**	0.20**
track genre: Jazz and Swing					−0.05	−0.05
track genre: Pop and Charts					−0.40***	−0.40***
track genre: Rock and Punk					0.05	0.05
track genre: Funk and Soul					−0.07*	−0.06*
residency: Germany (def: UK)						0.03
residency: Spain (def: UK)						0.12***
age group: generation X (def: gen. Y)						−0.03
age group: generation B (def: gen. Y)						0.08*
education: ISCED 3–4 (def: ISCED 0–2)						0.02
education: ISCED 5–8 (def: ISCED 0–2)						0.11***
gender: female (def: male)						0.07**
Nagelkerke’s $R^2$	18.57%	49.42%	50.71%	52.68%	52.92%	53.02%
Incremental $R^2$	18.57%	30.86%	1.29%	1.96%	0.24%	0.1%

Notes: all non-dummy predictors are standardized (beta-coefficients); standard errors are cluster-robust; track genres are effect-coded (redundant category: World Music); \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .



### 3. Study 2: Predicting Perceived Musical Expression by Algorithmic Audio Signal Analysis

In the course of Study 2 addressing RQ3 and RQ4, we developed computational prediction models explaining the non-individual parts of variance contained in the scores of perceived musical expression factors Arousal, Valence, Authenticity, Timeliness, and Eroticity (see Study 1).

#### 3.1. Methods

To this end, we utilised technical audio signal and music descriptors as predictors, which either stem from ML of branding experts' knowledge or algorithmic MIR toolboxes describing music and sound parameters. Details on development and selection of these predictor variables are provided in the Supplementary File (A2.1 and A2.2). Dependent variables were created by calculating the arithmetic mean of each perceived musical expression factor across all respondents of Study 1.

##### 3.1.1. Statistical Aggregation of Descriptors and Feature Selection for Computational Prediction Models

Linear hierarchical stepwise regression procedures were employed to aggregate the descriptors' explanatory power. In detail, predictor variables were always entered in a block-wise fashion, based on toolbox origin or ML descriptor group (see Table 11 in the Supplementary File, for a list of all predictor blocks). Within each block, a stepwise variable selection procedure (forward/backward-method with  $p_{in} = .05/p_{out} = .10$ ) was performed. We finally computed (incremental) adjusted  $R^2$  for each predictor block to estimate the explanatory power of the different descriptors.

#### 3.2. Results

##### 3.2.1. Accuracy of ML Classifiers

ML of the various classifiers led to very robust results (see Table 10 in the Supplementary File). ML Classification of *musical style* and the presence of *vocals* was accom-

plished with over 90% accuracy. By contrast, recognition of *instrumentation* (81% accuracy), *production timbre* (82% accuracy), and *vocals gender* (76% accuracy) turned out to be more challenging.

##### 3.2.2. Obtained Prediction Models

Across all computational models identified by hierarchical stepwise regression (see Table 11 in the Supplementary File), *musical style* and *instrumentation*, as learned by the ML algorithm play the most significant role in variance explanation of perceived musical expression ( $R_{adj}^2$  [style] = .191,  $R_{adj}^2$  [instrumentation] = .183). Also, rhythmic features extracted by the IRCAM beat toolbox explain a substantial amount of variance ( $R_{adj}^2$  [IRCAM beat] = .151), in particular related to perceived *Authenticity* and *Timeliness* of a musical excerpt. ( $R_{adj}^2$  [instrumentation] = .183). Finally, the remaining predictor blocks play a lesser important role in variance explanation, suggesting various interacting levels and facets of musical meaning. In the following, single prediction models obtained for the five musical expression factors will be described in detail.

##### 3.2.3. Valence

As already suggested by the overall results in Table 11 in the Supplementary File, *musical style* adherence probabilities ( $R_{adj}^2 = .177$ ) and instrumentation ( $R_{adj}^2 = .132$ ) play a crucial role in variance explanation of perceived Valence. Table 2 presents results of the hierarchical stepwise regression model, revealing *Hip Hop*, *Blues*, and *Oriental* as slightly associated with negative *Valence*, whereas *Samba*, *Rock and Roll*, and *Latin* are related to more positive *Valence*. Additionally, the probability of a track containing an *electric guitar* implies more negative *Valence*, possibly because electric guitars are often connotated as 'aggressive.' Finally, two production sound descriptors emerged in the list of the ten most potent predictors, namely the proportion of *noise energy* in the audio signal and its *periodicity* both being associated with more positive *Valence*.

**Table 2.** Hierarchical stepwise regression model predicting *Valence*, ten best predictors with largest  $\beta$  values.

Predictor	$\beta$	SE	t	p
Style (ML): HipHop	-0.178	0.035	-5.064	< .001
IRCAM descriptor: total noise energy	0.166	0.061	2.725	.007
IRCAM descriptor: periodicity	0.139	0.043	3.216	.001
Style (ML): Samba	0.139	0.035	3.967	< .001
Style (ML): Rock and Roll	0.136	0.033	4.058	< .001
Style (ML): Latin	0.129	0.035	3.712	< .001
IRCAM descriptor: sharpness SD	-0.127	0.044	-2.892	.004
Style (ML): Blues	-0.121	0.033	-3.709	< .001
Style (ML): Oriental	-0.119	0.032	-3.740	< .001
Intrumentation (ML): Electric Guitar	-0.118	0.035	-3.354	.001

### 3.2.4. Arousal

Adherence of an audio track to a *style* ( $R^2_{adj} = .239$ ) and its *instrumentation* ( $R^2_{adj} = .219$ ) also play a dominant role in variance explanation of *Arousal* (see Table 3). Musical *styles* such as *Downbeat*, *Balearic*, *Reggae*, *Boogie*, and *Soul* commonly associated with lower tempi and relaxation and calmness are major predictors of lowered arousal. The three best predictors, however, are directly related to production sound: Firstly, the more *harmonic energy* an audio track contains, the less it is perceived as arousing. This corroborates common knowledge in music psychology and psychoacoustics stating that the noisier (i.e., less harmonic) an audio track is, the more it is perceived as arousing (Juslin & Laukka, 2004). Secondly, the mean and standard deviation of the *first MFCC band* highlight the arousing role of the amount and fluctuation of low-frequency content (i.e., pumping beats) in a musical track. Finally, the model supports everyday experience that the more *percussive* and the less *warm* the sound of a musical track is, the more it will be perceived as arousing.

### 3.2.5. Authenticity

Regarding the attribution of *Authenticity*, *rhythmic* features as measured by the *IRCAM beat* toolbox ( $R^2_{adj} = .214$ ), as well as the adherence to a musical

*style* and associated image ( $R^2_{adj} = .169$ ), are crucial for variance explanation. Amongst the most critical features (Table 4), eight *styles* are negatively related to authenticity, four of them from the electronic dance music genre. This resonates with findings that the use of synthesised instruments, studio production, and more contemporary styles are often associated with lesser authenticity (Wu, Spieß, & Lehmann, 2017). The fact that instrumental (i.e., non-vocal) music, in general, predicts lesser *Authenticity* can be explained by assuming that it is foremost the vocal intonation of a singer that helps to represent human values such as being ‘honest.’ Finally, two production sound descriptors, namely *total harmonic energy* and fluctuations in the higher mid-frequency range (MFCC Band 05 [SD]) appear in the list. The former is associated with ‘non-distorted’ acoustical sounds in a track contributing to perceived *Authenticity*; the latter might be related to pulsating synthetic sounds occurring in electronic music and thus leading to less perceived authenticity.

### 3.2.6. Timeliness

Analogously to previous musical expression factors, musical *style* is also crucial for the variance explanation of *Timeliness* ( $R^2_{adj} = .213$ ), together with *instrumentation* ( $R^2_{adj} = .216$ ) and features related to *rhythm* ( $R^2_{adj}$  [IRCAM beat] = .297). Nine of the ten most potent single vari-

**Table 3.** Hierarchical stepwise regression model predicting *Arousal*, ten best predictors with largest  $\beta$  values.

Predictor	$\beta$	SE	t	p
IRCAM descriptor: total harmonic energy	-0.210	0.046	-4.531	< .001
MFCC Band 01 SD	0.193	0.037	5.205	< .001
MFCC Band 01 Mean	0.146	0.056	2.592	.010
Style (ML): Downbeat	-0.141	0.024	-5.856	< .001
IRCAM descriptor: Percussivity	0.124	0.031	4.029	< .001
Style (ML): Balearic	-0.114	0.023	-4.954	< .001
Production Timbre (ML): warm	-0.109	0.028	-3.831	< .001
Style (ML): Reggae	-0.100	0.023	-4.386	< .001
Style (ML): Boogie	-0.100	0.022	-4.476	< .001
Style (ML): Soul	-0.100	0.023	-4.416	< .001

**Table 4.** Hierarchical stepwise regression model predicting *Authenticity*, ten best predictors with largest  $\beta$  values.

Predictor	$\beta$	SE	t	p
Style (ML): UK Funky	-0.205	0.029	-7.125	< .001
Style (ML): Hip Hop	-0.203	0.035	-5.745	< .001
IRCAM descriptor: total harmonic energy	0.188	0.053	3.535	< .001
Vocals present (ML): no	-0.175	0.042	-4.148	< .001
Style (ML): Dubstep	-0.173	0.029	-6.000	< .001
Style (ML): Electro (ML)	-0.165	0.029	-5.635	< .001
MFCC Band 05 (SD)	-0.151	0.036	-4.136	< .001
Style (ML): Drum and Bass (ML)	-0.145	0.031	-4.740	< .001
Style (ML): Krautrock (ML)	-0.135	0.029	-4.714	< .001
Style (ML): Tech House (ML)	-0.134	0.031	-4.287	< .001

ables constitute *musical styles* that can be regarded as rather traditional (e.g., *German Schlager*, *Chanson*, *Classical Jazz*, *Country*) and were thus negatively associated with perceived Timeliness (Table 5). Also, the *proportion of noise* (i.e., non-harmonic) energy in an audio signal was a positive predictor of timeliness. High total noise energy often results from using (non-harmonic) synthetic sounds and effects, as typically found in rather modern industrial-sounding music styles (e.g., *Dubstep*).

### 3.2.7. Eroticity

Finally, concerning perceived *Eroticity* of a musical track, *instrumentation* ( $R^2_{adj} = .205$ ) and *style* ( $R^2_{adj} = .155$ ) explained the most substantial amount of variance (Table 6). A musical track is more likely to be perceived as erotic if containing female vocals, in particular as opposed to an instrumental track. In contrast, the presence of an *electric guitar* (presumably associated with rather ‘manly’ musical genres such as *Rock* and *Heavy Metal*) contributed negatively to perceived *Eroticity*. Moreover, *Soul* music is a positive predictor of *Eroticity*, whereas tracks from the styles *HipHop*, *Oriental*, and *UK Funky* styles are perceived as less erotic. Finally, a *warm* timbre as well as high mean values in the 11th MFCC band are related to stronger perceived *Eroticity* of a musical piece. The latter might be related to aspirated female vocals which are perceived as erotic.

### 3.3. Discussion

Findings from Study 2 demonstrate that it is possible to predict major portions of the non-individual parts of perceived expression in popular music with the aid of audio signal analysis, MIR, and ML techniques (RQ3). Inspecting obtained computational prediction models and addressing RQ4, it turned out that questions of *musical style*, *instrumentation*, and *rhythm* dominate perceived affective and semantic expressivity of popular music. Meanwhile, *production sound*, *keys*, *chords* and *lyrics* play only a minor role. Finally, we also found differences regarding the importance of specific musical elements when it came to different dimensions of musical expression. However, these were largely in line with existing research literature and too complex to be discussed here in further detail due to space limitations.

## 4. General Discussion

With the present contribution, we empirically compared different ways of explaining music liking in the ‘push scenarios’ which are becoming more prevalent in the age of digital media. Aiming at demonstrating the importance of the hitherto underestimated role of perceived musical expression, we compared its explanatory power with that of the received genre preference approach while controlling for well-known ‘cognitive side-effects’ in mu-

**Table 5.** Hierarchical stepwise regression model predicting *Timeliness*, ten best predictors with largest  $\beta$  values.

Predictor	$\beta$	SE	t	p
Style (ML): Schlager	-0.204	0.022	-9.270	< .001
Style (ML): Balkan	-0.202	0.022	-9.058	< .001
Style (ML): Oriental	-0.198	0.022	-9.139	< .001
Style (ML): Chanson	-0.195	0.023	-8.459	< .001
Style (ML): Asia	-0.176	0.022	-7.852	< .001
Style (ML): Calypso	-0.168	0.024	-6.941	< .001
Style (ML): Latin Style	-0.166	0.024	-6.984	< .001
Style (ML): Classical Jazz	-0.165	0.025	-6.606	< .001
IRCAM descriptor: Total Noise Energy	0.158	0.039	4.042	< .001
Style (ML): Country	-0.155	0.023	-6.618	< .001

**Table 6.** Hierarchical stepwise regression model predicting *Eroticity*, ten best predictors with largest  $\beta$  values.

Predictor	$\beta$	SE	t	p
Vocals (ML): no	-0.221	0.039	-5.617	< .001
Female vocals (ML): yes	0.221	0.044	4.987	< .001
Style (ML): HipHop	-0.156	0.038	-4.082	< .001
Instrumentation (ML): Electric Guitar	-0.149	0.034	-4.371	< .001
Style (ML): Oriental	-0.137	0.031	-4.409	< .001
Production timbre (ML): Dark	-0.135	0.034	-3.958	< .001
MFCC Band 11 MEAN	0.130	0.034	3.892	< .001
Style (ML): Soul	0.127	0.032	4.015	< .001
IRCAM key: Db (effect-coded)	0.121	0.165	3.368	0.001
Style (ML): UK Funky	-0.116	0.031	-3.686	< .001

music liking (Study 1). While the latter (foremost familiarity, but also prominence and popularity) were shown to explain a fair amount of variance, the explanatory potential of genre affinities expectedly turned out to be minor compared to the influence of perceived musical expression. Notably, advancing the state of research in the field, we demonstrated that perceived semantic meaning is as important as perceived strength of expressed emotions when it comes to explaining liking for previously unknown music. In summary, all our hypotheses were confirmed. Additionally, attribution of meaning towards presented and largely previously unknown music was found to be particularly homogenous across sociodemographic groups and countries. Similarly, sociodemographic differences regarding the weighting of different musical expression dimensions for music liking turned out to be small. Nevertheless, a significant degree of individual (presumably also encompassing situational) heterogeneity in musical meaning attribution still exists.

Based on these findings, we used MIR and ML in Study 2 to test the algorithmic predictability of the perceived musical expression. As expected, it turned out that meaning attribution concerning popular music appears to a substantial degree to be uniform and rule-based. The explanatory power of musical style in Study 2, when compared to our findings regarding the related, but coarser concept of musical genre in Study 1 hints at the possibility that fine-grained, highly standardised algorithmic style descriptors instead of subjective ratings might form a solution for the ‘genre dilemma’ discussed in the introduction and the research literature (Brisson & Bianchi, 2019). Taken together, the findings point out the importance of *communicative aspects of popular music* when it comes to empirically explaining and predicting music liking in basic musicological research on music preferences as well as in applied scenarios such as music recommendation algorithms.

Overall, the findings of our two studies stress the importance of a hitherto underdeveloped area in quantitative music reception research: music semantics. Previously, music psychology tended to analyse popular music predominantly as an art form or as a sensual media offering that may emotionally move us and entrain our bodies into dancing. However, with this contribution, we suggest conceiving of popular music also as a semiotic device, a carrier of complex meanings, similar to oral language or any other communicative sign system. This can be interpreted in terms of music’s anthropological main functions of self-awareness and social relatedness (Schäfer, Sedlmeier, Städtler, & Huron, 2013). Popular music once more presents itself as something that brings people together, not only in terms of affect, but also in terms of identity and values (Frith, 1996). Until now, however, expression of these aspects in pop music have been researched predominantly by cultural studies scholars, either by employing discourse analysis (Machin & Richardson, 2012) or interpretive interview studies (Hesmondhalgh, 2007). Here, our paper demonstrates

that meaning structures in music exert strong measurable quantitative effects, and that these are relatively homogenous across social groups and cultures, making them well-suited for statistical analyses with larger samples and also largely predictable by ML.

Several limitations regarding the generalisation of our findings have to be addressed. Popular music is a complex global cultural phenomenon, and the existing repertoire of genres, styles, artists and scenes is vast. Our study was only able to analyse music listeners from three European countries and only employed a very limited, though comparatively heterogeneous selection of popular music. In general, it appears hard to claim with any sample of any size to have a proper representation of popular music as such, due to its breadth, complexity, and everchanging nature. Furthermore, we conducted a secondary analysis of popular music titles all deemed suitable for branding purposes, which necessarily leads to the exclusion of more extreme, fringe styles of pop music. The finding that the ML features operationalising the content of song lyrics did not play a substantial role in the final models of Study 2 could be related to this fact. Further, it is crucial to acknowledge that—by design—the musical expression space operationalised by the GMBI\_15 questionnaire instrument does not exhaust the full breadth of musical expression. Hence, further research should expand from our findings, especially with a sharper focus on the expression of identity and humanistic, political and religious values.

Summarising implications, we propose that musicology should consider taking a shift in research focus ‘from mood to meaning’ (Vorderer & Reinecke, 2015) that has already taken place in media research. The observed importance of the authenticity dimension further parallels the claim of a ‘truth-seeking media recipient’ that has recently gained prominence in media entertainment research (Oliver & Raney, 2011). Also, our findings suggest that popular music’s meaning expression is a legitimate field of research for applying communication theory, because it appears to act like a rule-based language, as demonstrated by Study 2. It can thus be analysed similarly to linguistic or pictorial content and may also form an independent variable in media reception and effects research (Shevy, 2013). In conclusion, we argue that the results of our studies are of importance not only for the music industry and musicology but also for media studies and communication science.

### Acknowledgments

This work was supported by the European Union’s Horizon 2020 research and innovation program under grant agreement No 688122. We express our gratitude to Geoffroy Peeters and his IRCAM team for their contributions to the ML parts of Study 2. Similarly, we thank our project partner HearDis for contributing the stimulus material and tag knowledge, as well as Dr Andreas Schönrock for stimulus pool preparation.

## Conflict of Interests

The authors declare no conflict of interests.

## Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

## References

- Airoidi, M., Beraldo, D., & Gandini, A. (2016). Follow the algorithm: An exploratory investigation of music on YouTube. *Poetics*, 57, 1–13. <https://doi.org/10.1016/j.poetic.2016.05.001>
- Berlyne, D. E. (1971). *Aesthetics and psychobiology*. New York, NY: Appleton Century-Crofts.
- Bourdieu, P. (1984). *Distinction*. Polity Press.
- Bouvier, G., & Machin, D. (2013). How advertisers use sound and music to communicate ideas, attitudes and identities: A multimodal critical discourse approach. In B. Pennock-Spek & D. S. Rubio (Eds.), *The multimodal analysis of television commercials*. Valencia: University of Valencia Press.
- Brisson, R., & Bianchi, R. (2019). On the relevance of music genre-based analysis in research on musical tastes. *Psychology of Music*. Advance online publication. <https://doi.org/10.1177/0305735619828810>
- Brunswik, E. (1952). *The conceptual framework of psychology*. Illinois, IL: University of Chicago Press.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233–255. [https://doi.org/10.1207/S15328007SEM0902\\_5](https://doi.org/10.1207/S15328007SEM0902_5)
- Chmiel, A., & Schubert, E. (2017). Back to the inverted-U for music preference: A review of the literature. *Psychology of Music*, 45(6), 886–909. <https://doi.org/10.1177/0305735617697507>
- Drott, E. (2018). Why the next song matters: Streaming, recommendation, scarcity. *Twentieth-Century Music*, 15(3), 325–357. <https://doi.org/10.1017/S1478572218000245>
- Eerola, T., Friberg, A., & Bresin, R. (2013). Emotional expression in music: Contribution, linearity, and additivity of primary musical cues. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00487>
- Egermann, H., Grewe, O., Kopiez, R., & Altenmüller, E. (2009). Social feedback influences musically induced emotions. *Annals of the New York Academy of Sciences*, 1169(1), 346–350. <https://doi.org/10.1111/j.1749-6632.2009.04789.x>
- Feezell, J. T. (2017). It's not only rock and roll: The influence of music preferences on political attitudes. In *Music as a platform for political communication* (pp. 167–186). Hershey, PA: IGI Global.
- Fricke, K. R., & Herzberg, P. Y. (2017). Personality and self-reported preference for music genres and attributes in a German-speaking sample. *Journal of Research in Personality*, 68, 114–123. <https://doi.org/10.1016/j.jrp.2017.01.001>
- Frith, S. (1996). *Performing rites: On the value of popular music*. Oxford: Oxford University Press.
- Greenberg, D. M., Kosinski, M., Stillwell, D. J., Monteiro, B. L., Levitin, D. J., & Rentfrow, P. J. (2016). The song is you: Preferences for musical attribute dimensions reflect personality. *Social Psychological and Personality Science*, 7(6), 597–605. <https://doi.org/10.1177/1948550616641473>
- Hennion, A. (2001). Music lovers: Taste as performance. *Theory Culture Society*, 18(5), 1–22. <https://doi.org/10.1177/02632760122051940>
- Herzog, M., Lepa, S., & Egermann, H. (2016). *Towards automatic music recommendation for audio branding scenarios*. Paper presented at 17th International Society for Music Information Retrieval Conference (ISMIR), New York, NY.
- Herzog, M., Lepa, S., Egermann, H., Schönrock, A., & Steffens, J. (2020). Towards a common terminology for music branding campaigns. *Journal of Marketing Management*, 36(1/2), 176–209. <https://doi.org/10.1080/0267257X.2020.1713856>
- Herzog, M., Lepa, S., Steffens, J., Schönrock, A., & Egermann, H. (2017a). Predicting musical meaning in audio branding scenarios. In E. Van Dyck (Ed.), *Proceedings of the 25th Anniversary Conference of the European Society for the Cognitive Sciences of Music (ESCOM)*. Ghent: ESCOM.
- Herzog, M., Lepa, S., Steffens, J., Schönrock, A., & Egermann, H. (2017b). *Predicting musical meaning in audio branding scenarios*. Paper presented at Conference of the European Society for the Cognitive Sciences of Music (ESCOM), 2017, Ghent.
- Hesmondhalgh, D. (2007). Audiences and everyday aesthetics. Talking about good and bad music. *European Journal of Cultural Studies*, 10(4), 507–527. <https://doi.org/10.1177/1367549407081959>
- Hervner, K. (1936). Experimental studies of the elements of expression in music. *The American Journal of Psychology*, 48(2), 246–268.
- Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, 110(3), 306–340.
- Juslin, P. N. (2000). Cue utilization in communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance*, 26(6), 1797–1813.
- Juslin, P. N., & Laukka, P. (2004). Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening. *Journal of New Music Research*, 33(3), 217. <https://doi.org/10.1080/0929821042000317813>
- Klimmt, C. (2011). Media psychology and complex modes of entertainment experiences. *Journal of*



- Media Psychology: Theories, Methods, and Applications*, 23(1), 34–38. <https://doi.org/10.1027/1864-1105/a000030>
- Krämer, B. (2018). Online music recommendation platforms as representations of ontologies of musical taste. *Communications*, 43(2), 259–281. <https://doi.org/10.1515/commun-2017-0056>
- Kristen, S., & Shevy, M. (2013). A comparison of German and American listeners' extra musical associations with popular music genres. *Psychology of Music*, 41(6), 764–778. <https://doi.org/10.1177/0305735612451785>
- Lahire, B. (2008). The individual and the mixing of genres: Cultural dissonance and self-distinction. *Poetics*, 36(2/3), 166–188. <https://doi.org/10.1016/j.poetic.2008.02.001>
- Lepa, S. (2010). Media appropriation from a critical realist point of view? Theoretical premises and empirical research. In A. S. da Silva, J. C. Martins, L. Magalhães, & M. Goncalves (Eds.), *Comunicação, cognição e mídia. Volume 1* (pp. 379–390). Lisbon: Aletheia.
- Lepa, S., & Hoklas, A.-K. (2015). How do people really listen to music today? Conventionalities and major turnovers in German audio repertoires. *Information, Communication & Society*, 18(10), 1253–1268. <https://doi.org/10.1080/1369118X.2015.1037327>
- Lonsdale, A. J., & North, A. C. (2009). Musical taste and ingroup favouritism. *Group Processes & Intergroup Relations*, 12(3), 319–327. <https://doi.org/10.1177/1368430209102842>
- Lonsdale, A. J., & North, A. C. (2012). Musical taste and the representativeness heuristic. *Psychology of Music*, 40(2), 131–142. <https://doi.org/10.1177/0305735611425901>
- Machin, D., & Richardson, J. E. (2012). Discourses of unity and purpose in the sounds of fascist music: A multimodal approach. *Critical Discourse Studies*, 9(4), 329–345. <https://doi.org/10.1080/17405904.2012.713203>
- Madison, G., & Schiölde, G. (2017). Repeated listening increases the liking for music regardless of its complexity: Implications for the appreciation and aesthetics of music. *Frontiers in Neuroscience*, 11. <https://doi.org/10.3389/fnins.2017.00147>
- Mellander, C., Florida, R., Rentfrow, P. J., & Potter, J. (2018). The geography of music preferences. *Journal of Cultural Economics*, 42(4), 593–618. <https://doi.org/10.1007/s10824-018-9320-x>
- Oliver, M. B., & Raney, A. A. (2011). Entertainment as pleasurable and meaningful: Identifying hedonic and eudaimonic Motivations for entertainment consumption. *Journal of Communication*, 61(5), 984–1004. <https://doi.org/10.1111/j.1460-2466.2011.01585.x>
- Peterson, R. A. (1992). Understanding audience segmentation: From elite and mass to omnivore and univore. *Poetics*, 21(4), 243–258. [https://doi.org/10.1016/0304-422X\(92\)90008-Q](https://doi.org/10.1016/0304-422X(92)90008-Q)
- Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 123–205). Cambridge, MA: Academic Press.
- Pontara, T., & Volgsten, U. (2017). Musicalization and mediatization. In O. Driessens, G. Bolin, A. Hepp, & S. Hjarvard (Eds.), *Dynamics of mediatization* (pp. 247–269). New York, NY: Springer International Publishing. [https://doi.org/10.1007/978-3-319-62983-4\\_12](https://doi.org/10.1007/978-3-319-62983-4_12)
- Prior, N. (2013). Bourdieu and the sociology of music consumption: A critical assessment of recent developments. *Sociology Compass*, 7(3), 181–193. <https://doi.org/10.1111/soc4.12020>
- Rentfrow, P. J., Goldberg, L. R., & Levitin, D. J. (2011). The structure of musical preferences: A five-factor model. *Journal of Personality and Social Psychology*, 100(6), 1139–1157. <https://doi.org/10.1037/a0022406>
- Rentfrow, P. J., Goldberg, L. R., Stillwell, D. J., Kosinski, M., Gosling, S. D., & Levitin, D. J. (2012). The song remains the same: A replication and extension of the music model. *Music Perception: An Interdisciplinary Journal*, 30(2), 161–185. <https://doi.org/10.1525/mp.2012.30.2.161>
- Roose, H., & Stichele, A. V. (2010). Living room vs. concert hall: Patterns of music consumption in Flanders. *Social Forces*, 89(1), 185–207. <https://doi.org/10.1353/sof.2010.0077>
- Schäfer, T., & Mehlhorn, C. (2017). Can personality traits predict musical style preferences? A meta-analysis. *Personality and Individual Differences*, 116, 265–273. <https://doi.org/10.1016/j.paid.2017.04.061>
- Schäfer, T., Sedlmeier, P., Städtler, C., & Huron, D. (2013). The psychological functions of music listening. *Frontiers in Psychology*, 4(511). <https://doi.org/10.3389/fpsyg.2013.00511>
- Searle, S. R. (1995). An overview of variance component estimation. *Metrika*, 42(1), 215–230. <https://doi.org/10.1007/BF01894301>
- Shevy, M. (2008). Music genre as cognitive schema: Extramusical associations with country and hip-hop music. *Psychology of Music*, 36(4), 477–498. <https://doi.org/10.1177/0305735608089384>
- Shevy, M. (2013). Integrating media effects research and music psychology. In S.-L. Tan, A. J. Cohen, S. D. Lipscomb, & R. A. Kendall (Eds.), *The psychology of music in multimedia* (pp. 66–88). Oxford: Oxford University Press.
- Sievers, B., Polansky, L., Casey, M., & Wheatley, T. (2012). Music and movement share a dynamic structure that supports universal expressions of emotion. *Proceedings of the National Academy of Sciences*, 110(1), 70–75. <https://doi.org/10.1073/pnas.1209023110>
- Stone, A. (2016). Meaning and affect in popular music. In A. Stone (Ed.), *The value of popular music: An approach from post-kantian aesthetics* (pp. 173–211). New York, NY: Springer International Publishing. [https://doi.org/10.1007/978-3-319-46544-9\\_6](https://doi.org/10.1007/978-3-319-46544-9_6)
- Tagg, P. (2000). Folk music, art music, popular music: An



axiomatic triangle. In *Kojak. Fifty seconds of television music. Towards the analysis of affect in popular music* (2nd ed., p. 35). Larchmont, NY: Mass Media Music Scholars' Press.

Tagg, P. (2013). *Music's meanings: A modern musicology for non-musos*. Larchmont, NY: Mass Media Music Scholar's Press.

Vlegels, J., & Lievens, J. (2017). Music classification, genres, and taste patterns: A ground-up network analysis on the clustering of artist preferences. *Poetics*, 60, 76–89. <https://doi.org/10.1016/j.poetic.2016.08>.

004

Vorderer, P., & Reinecke, L. (2015). From mood to meaning: The changing model of the user in entertainment research. *Communication Theory*, 25(4), 447–453. <https://doi.org/10.1111/comt.12082>

Wu, L., Spieß, M., & Lehmann, M. (2017). The effect of authenticity in music on the subjective theories and aesthetical evaluation of listeners: A randomized experiment. *Musicae Scientiae*, 21(4), 1–23. <https://doi.org/10.1177/1029864916676301>

## About the Authors



**Steffen Lepa** (PhD), born 1978, is a Postdoc Researcher and Lecturer at Audio Communication Group, TU Berlin. He holds MA studies in media, psychology, computer science and communication science and a PhD in educational and social sciences. He teaches on digital media change, research methodology, and sound design. In 2018–2019 he was a Visiting Professor for Media and Music at Hanover University of Music, Drama and Media. His current key research areas are mediatization, audio reception, audience studies (focus on music streaming and analogue media) and computational research methods.



**Jochen Steffens** (PhD) born 1981, is a music and sound Researcher and a Professor of Musical Acoustics at Düsseldorf University of Applied Sciences. After completing his PhD, he held postdoctoral positions at McGill University, the Max Planck Institute for Empirical Aesthetics and the TU Berlin, where he was awarded his Habilitation in systematic musicology and psychoacoustics. His research areas include the perception of music and sound, their effects on emotional, cognitive and behavioural processes, and associated practical applications.



**Martin Herzog** graduated in Computer Science at Humboldt University Berlin in 2011. Since 2016, he is a Research Associate at the Audio Communication Group at TU Berlin. His research areas include music perception, music information retrieval, and audio branding. In his PhD research he focusses on predicting musical meaning from high-level music features. Apart from science, Martin Herzog works as a consultant for digital communications in Berlin since 2011.



**Hauke Egermann** (PhD) is an Associate Professor in the Department of Music, University of York. He graduated in Systematic Musicology, Media Studies, and Communication Research (MA 2006, Hanover University for Music, Drama and Media). Subsequently, he studied at the Centre for Systems Neurosciences Hanover (PhD, 2009). He was Postdoctoral Research Fellow at the Centre for Interdisciplinary Research in Music Media and Technology (2009–2011, McGill University, Montreal, Canada). In 2016, he was awarded his Habilitation in Musicology at the Technische Universität Berlin. He directs the York Music Psychology Group since 2016.