scientific reports

OPEN

Music communicates social emotions: Evidence from 750 music excerpts

Elliot X. Pring[1], Kirk N. Olsen[2], Anthony E. D. Mobbs[1] & William Forde Thompson[1,3]✉

Humans perceive a range of basic emotional connotations from music, such as joy, sadness, and fear, which can be decoded from structural characteristics of music, such as rhythm, harmony, and timbre. However, despite theory and evidence that music has multiple social functions, little research has examined whether music conveys emotions specifically associated with social status and social connection. This investigation aimed to determine whether the social emotions of *dominance* and *affiliation* are perceived in music and whether structural features of music predict social emotions, just as they predict basic emotions. Participants (*N* = 1513) listened to subsets of 750 music excerpts and provided ratings of energy arousal, tension arousal, valence, dominance, and affiliation. Ratings were modelled based on ten structural features of music. Dominance and affiliation were readily perceived in music and predicted by structural features including rhythm, harmony, dynamics, and timbre. In turn, energy arousal, tension arousal and valence were also predicted by musical structure. We discuss the results in view of current models of music and emotion and propose research to illuminate the significance of social emotions in music.

Keywords Basic emotions, Music perception, Social emotions, Dominance, Affiliation

Acoustic features of music communicate basic and social emotions: Evidence from 750 music excerpts

Music has the capacity to express and evoke a wide range of emotions[1–3]. Listeners can identify emotional connotations after only a few seconds of music ("perceived emotion"), whereas it takes longer to induce an emotional experience ("felt emotion")[4–6]. In this investigation, we describe the results of an international crowd-sourcing survey in which 1563 participants listened to different subsets of 750 music excerpts and provided ratings of perceived emotions. Regression and principal-component analyses were conducted to identify optimal models of emotional responses elicited by musical attributes, including models of two *social emotions* that have not been investigated with respect to music.

Emotional responses to music are typically assessed within the frameworks of discrete or dimensional models. Discrete models employ emotion labels such as happiness, sadness, anger, and disgust, which are sometimes grouped into types such as social emotions (e.g., feelings of connection, empowerment[7]); moral emotions (e.g., remorse, righteous indignation[8]); aesthetic emotions (e.g., awe, transcendence); achievement-related emotions (e.g., pride, disappointment[9]); and epistemic emotions (e.g., curiosity, doubt[10]). Such groupings acknowledge that emotions are tethered to the causal and contextual circumstances associated with the feeling state[11].

Dimensional models depict emotions as points on underlying affective continua, such as the degree of energy experienced in the emotion[12]. Dimensional models capture complex emotions that vary in intensity but are difficult to label[13]. For example, death metal fans may perceive complex mixtures of tension, empowerment, energy, and joy in their preferred music, making it difficult to label[14,15]. The circumplex model maps emotion onto two dimensions of affect: arousal and valence[16–18]. In some models, arousal is further divided into two forms of activation: *energy* arousal (EA), which ranges from tired to energetic, and *tension* arousal (TA), which ranges from calm to nervous[19,20].

The emotions expressed by music are encoded in both compositional structure and performance attributes[21,60], including mode, tempo, dynamics (loudness), pitch register, articulation, and timbre[22–24]. For

[1]School of Psychological Sciences, Macquarie University, Sydney, Australia. [2]Australian Institute of Health Innovation, Macquarie University, Sydney, Australia. [3]Faculty of Society and Design, Bond University, Gold Coast, Australia. ✉email: bthompso@bond.edu.au

example, fast tempo, consonance, major mode, and bright timbre are associated with happiness, high arousal, and positive valence[21,25].

Research on music and emotion has predominantly focused on basic emotions such as joy, sadness, fear, and anger[2]. These emotions reflect abstract feeling states that may be experienced by an individual across varying contexts. However, music is often experienced in social settings, including cultural rituals, music festivals, and ceremonies, and the social functions of music are emphasised across disciplines. Music can trigger a sense of social inclusion for fans (social affiliation) or social exclusion for non-fans[2,14]. Music can also convey feelings of social status, ranging from empowerment (dominance) to submissiveness and disempowerment[11,15].

Dominance and affiliation are core social emotions[26–29], and represent two axes of a social emotion space[28]. Dominance encompasses feelings ranging from leadership and empowerment over others to subordination and powerlessness, while affiliation refers to feelings that range from social connectedness to feelings of isolation and loneliness. These concepts can sometimes reflect social judgments or dispositions, but they also describe feeling states[26–29].

Recent research further justifies examination of these two social emotions. Aucouturier and Canonne[38] proposed that music is not only social in its production but also in its perception, as it can be perceived as the sonic trace of social relations between agents. Their investigation showed that musically trained and untrained listeners could decode social intentions, such as dominance and affiliation, from the acoustic features of improvised musical interactions. Other research highlights the broader impact of music on social cognition, whereby exposure to certain musical elements in music from other cultures can influence listeners' evaluations of those cultural groups[59]. Finally, functional neuroimaging studies have revealed distinct neural pathways for dominance and affiliation[30], and dominance has been proposed as a core dimension of affect within the circumplex model[31–34].

The current research builds on this foundation by examining the capacity for music to communicate social emotions in 750 musical samples across multiple genres with responses from over 1500 listeners. Bipolar scales measuring dominance and affiliation[28] were used to assess the degree to which these emotions are perceived. Music analysis software and statistical modelling were used to model how various musical attributes predict the emotional meaning perceived by listeners, including the social emotions of dominance and affiliation and the emotional dimensions of EA, TA, and valence. The study also examined whether musical attributes predict responses to dimensional models of emotion. Musical stimuli were restricted to 5-second excerpts to enable the recruitment of a large sample of participants rating multiple music excerpts in a single testing session.

Sixteen structural elements such as rhythm, dynamics, harmony, and timbre were extracted from samples[34–36]. Modelling determined whether structural elements predict perceived EA, TA, valence, dominance, and affiliation in music. Several a priori hypotheses were established. First, we predicted that social emotions (dominance and affiliation) and perceived EA, TA, and valence should be perceived in music samples, as reflected in mean ratings (H1). Second, we anticipated that social emotions and perceived EA, TA, and valence should be predicted by structural elements such as rhythm, timbre, and pitch height (H2)[19,37].

Results

Descriptive statistics
Descriptive statistics are shown in Table 1, with responses to the Modified Self-Assessment Manikin (SAM) scale converted to a value between 1 and 7. Given the large sample sizes, the normality of sampling distributions was assumed for all variables. All means in Table 1 are reliably higher than a neutral rating of 4.0 ($p < .0001$), which would be expected if the emotion was not perceived. Standard deviations indicate the typical difference between the mean rating assigned to individual music samples and the mean rating averaged across all 750 music samples.

Correlation between emotion ratings
An examination of correlations between the various emotion ratings (See Table 2) revealed a moderate and significant positive correlation between EA and TA, and a moderate and significant negative correlation between TA and valence. Ratings of dominance were highly correlated with ratings of energy arousal and tension arousal, suggesting that this social emotion might be influenced by arousal levels perceived in music excerpts.

Correlations between music attributes and emotion ratings
Bivariate correlations were calculated between the attributes of music excerpts and perceived emotion ratings of EA, TA, valence, dominance, and affiliation (see Table 3). There were significant correlations between several music attributes and measures of EA, TA, dominance, affiliation, which were mainly moderate in strength. Correlations between valence and music attributes were weak or non-significant. Tempo was highly

Variable	N	M	SD	Skewness	Kurtosis
EA	749	5.06	0.69	−0.30	2.08
TA	749	4.26	0.70	−0.03	2.32
Valence	749	5.05	0.44	−0.47	2.94
Dominance	749	4.69	0.56	−0.12	2.40
Affiliation	749	4.84	0.44	−0.02	2.84

Table 1. Means, standard deviations, skewness, and kurtosis of dependent variables. *Note: N* = music stimuli; EA = energy arousal; TA = tension arousal.

Variable	EA	TA	Valence	Dominance	Affiliation
EA	-				
TA	0.66**	-			
Valence	−0.08*	−0.56**	-		
Dominance	0.80**	0.82**	−0.29**	-	
Affiliation	0.62**	0.12**	0.44**	0.39**	-

Table 2. Correlations between dependent variables. *$p < .05$ **$p < .01$.

Variable	EA	TA	Valence	Dominance	Affiliation
Tempo	0.90**	0.72**	−0.17**	0.81**	0.55**
Pulse clarity	0.46**	0.20**	0.06	0.28**	0.40**
RMSE	0.18**	0.12**	−0.05	0.28**	0.40**
Brightness	0.63**	0.46**	−0.13**	0.53**	0.36**
Dissonance	0.23**	0.30**	−0.21**	0.29**	0.05
Mode	0.02	−0.07*	0.09*	−0.02	0.08*
Spectral centroid	0.59**	0.37**	−0.04	0.45**	0.42**
Spectral rolloff	0.56**	0.32**	−0.00	0.42**	0.43**
Mean pitch	0.26*	0.20**	−0.07	0.26**	0.14*
Event density	0.39**	0.36**	−0.18**	0.36**	0.15**
Spectral spread	0.52**	0.25**	0.05	0.37**	0.44**
Spectral skewness	−0.53**	−0.34**	0.05	−0.44**	−0.38**
Spectral kurtosis	−0.32**	−0.21**	0.00	−0.27**	−0.25**
Spectral flatness	0.40**	0.17**	0.07	0.27**	0.35**
Spectral entropy	0.68**	0.48**	−0.12**	0.57**	0.43**
ZCR	0.53**	0.42**	−0.12**	0.45**	0.28**

Table 3. Correlations of psychophysical music attributes and ratings scales. *Note:* Psychophysical music attributes are on the left of the table and the emotions that were rated are listed at the top of each column. EA = energy arousal; TA = tension arousal; RMSE = root mean square energy; ZCR = zero-crossing rate. *$p < .05$. **$p < .01$.

correlated with EA ($r = .90$), TA ($r = .72$), and dominance ($r = .81$). Brightness, spectral centroid, spectral rolloff, event density, spectral entropy, and zero-crossing rate returned moderate positive correlations with EA, TA, dominance, and affiliation.

Music feature reduction: addressing collinearity with principal component analysis

Principal Component Analysis (PCA) was used to address multicollinearity among the extracted music features, a known issue due to overlapping computational methods in the MIRtoolbox[47]. PCA identified communalities among psychophysical attributes, allowing us to retain only the most significant factors while ensuring that regression models remained parsimonious. As shown in Table 4, the first component was dominated by spectral features, leading us to retain only spectral centroid, brightness, and zero-crossing rate. The second component included both RMSE (loudness) and dissonance, which were retained due to their distinct computational processes and relevance to perceived tension. The third component, involving pulse clarity and event density, also led to retaining both features, as they reflect different aspects of rhythmic clarity. Finally, mode was retained from the fourth component due to its significant loading and relevance to valence perception. Overall, six music features were excluded, resulting in the retention of ten features across five categories: timbre (ZCR, spectral centroid and brightness); harmony (dissonance and mode); dynamics (RMSE); rhythm (tempo, pulse clarity, and event density), and pitch.

Models of music features and perceived emotional dimensions

Regression model: Energy arousal

There were 749 observations (music stimuli scored according to their musical attributes and perceived affect). The Shapiro-Wilk test indicated that residuals were normally distributed ($p > 0.05$), and inspection of the plot of residuals against fitted values indicates the assumption of linearity was met. However, there was noticeable clustering towards the centre of the residuals plot. White's heteroscedasticity test indicated that the homoscedasticity assumption was not violated ($p = .129$). Therefore, inference testing was conducted with a multiple least-squares regression model.

The model assessing whether musical attributes are associated with perceived expression of energy arousal in music was significant, $F(10, 738) = 342.66$, $p < .001$, and explained 82.28% of the variance in the perception

Variable	1	2	3	4
Pulse clarity	0.36	0.28	**0.52**	0.18
RMSE	0.17	**0.86**	−0.27	−0.03
Brightness	**0.92**	−0.14	0.10	−0.06
Dissonance	0.20	**0.82**	−0.10	−0.08
Mode	0.10	−0.06	−0.00	**0.85**
Spectral centroid	**0.97**	−0.07	0.07	−0.03
Spectral rolloff	**0.95**	−0.01	0.00	−0.01
Mean pitch	0.31	0.05	0.26	−0.44
Event density	0.23	0.49	**0.49**	0.20
Spectral spread	**0.92**	0.05	−0.12	0.01
Spectral skewness	−**0.88**	0.05	0.30	−0.04
Spectral kurtosis	−**0.59**	0.04	**0.50**	−0.09
Spectral flatness	**0.80**	−0.06	−0.12	0.04
Spectral entropy	**0.94**	−0.08	0.04	−0.04
ZCR	**0.72**	−0.26	0.37	−0.06

Table 4. Component loadings for musical attributes. *Note:* The highest loadings for each component have been emboldened.

Variable	Coef.	SE	t	*p*	95% CI
Tempo	0.78	0.23	34.04	**< 0.001**	[0.74, 0.83]
Pulse clarity	0.19	0.06	3.02	**0.003**	[0.07, 0.32]
RMSE	-0.16	0.57	-0.28	0.782	[-1.28, 0.96]
Spectral centroid	-0.00	0.00	-0.40	0.686	[0.00, 0.00]
Brightness	0.54	0.17	3.15	**0.002**	[0.20, 0.88]
Dissonance	0.00	0.00	0.58	0.565	[0.00, 0.00]
Mode	0.13	0.08	1.74	0.082	[-0.02, 0.01]
Event density	-0.01	0.01	-1.24	0.217	[-0.03, 0.01]
Mean pitch	0.00	0.00	0.64	0.522	[0.00, 0.00]
ZCR	0.00	0.00	-1.29	0.197	[0.00, 0.00]

Table 5. Regression statistics for energy arousal model.

of EA in music (unadjusted estimate). There were three statistically significant relationships observed: EA and tempo ($\eta_p^2 = 0.636$), EA and pulse clarity ($\eta_p^2 = 0.008$), and EA and brightness ($\eta_p^2 = 0.013$) (see Table 5). These relationships were all positive, indicating that as tempo, pulse clarity, or brightness increase, so does the relative perceived magnitude of EA in music when all other variables in the model are held constant. While pulse clarity and brightness were highly significant predictors of perceptions of EA in music, the effect sizes of both predictors were small. EA is closely linked to rhythmic and timbral elements, with the focal point being pulse tempo.

Regression model: Tension arousal
The Shapiro-Wilk test confirmed the assumption of normality of residuals was met ($p > .05$). The assumption of linearity was also met after inspection of the residuals and fitted values plot. Slight clustering towards the centre of the residuals against the plot of the fitted values was observed; however, White's test indicated the assumption of homoscedasticity was not violated ($p = .381$). Therefore, multiple regression was deemed appropriate.

The general model determining if musical attributes are associated with the perception of TA in music was significant, $F(10, 738) = 102.01$, $p < .001$, explaining 58.02% of the variance in perceptions of TA in music. Table 6 illustrates four significant negative relationships in the model: TA and RMSE ($\eta_p^2 = 0.005$), TA and pulse clarity ($\eta_p^2 = 0.027$), TA and spectral centroid ($\eta_p^2 = 0.022$), and TA and mode ($\eta_p^2 = 0.006$). These relationships indicate that the perception of TA in music decreases in magnitude if the pulse of the music is more discernible or if there is more overall energy in the excerpt. Moreover, perceived TA in music decreases in magnitude when the spectral centroid of mass frequency increases or when the modality value is more positive, holding all other musical attributes constant.

There were three significant positive relationships in this model: TA and tempo ($\eta_p^2 = 0.344$), TA and brightness ($\eta_p^2 = 0.011$), and TA and dissonance ($\eta_p^2 = 0.032$) (see Table 6). These relationships indicate that the perception of TA in music increases in magnitude when either the speed of the musical beat increases, the amount of high-frequency energy in the music increases, or the amount of dissonance present in the music excerpt increases. These positive and negative relationships indicate that perceptions of TA in music are influenced by rhythmic (tempo and pulse clarity), harmonic (mode and dissonance), timbral (spectral centroid

Variable	Coef.	SE	t	p	95% CI
Tempo	0.71	0.36	19.67	**<0.001**	[0.64, 0.78]
Pulse clarity	-0.46	0.10	-4.56	**<0.001**	[-0.65, -0.26]
RMSE	-1.78	0.89	-2.00	**0.046**	[-3.52, -0.03]
Spectral centroid	0.00	0.00	-4.09	**<0.001**	[0.00, 0.00]
Brightness	0.78	0.27	2.90	**0.004**	[0.25, 1.31]
Dissonance	0.00	0.00	5.00	**<0.001**	[0.00, 0.00]
Mode	-0.26	0.12	-2.13	**.033**	[-0.49, -0.02]
Event density	0.01	0.01	1.09	0.277	[-0.01, 0.04]
Mean pitch	0.00	0.00	-0.13	0.896	[0.00, 0.00]
ZCR	0.00	0.00	1.66	0.098	[0.00, 0.00]

Table 6. Regression statistics for tension arousal model.

Variable	Coef.	SE	t	p	95% CI
Tempo	-0.06	0.03	-1.91	0.057	[-0.13, 0.00]
Pulse clarity	0.29	0.09	3.20	**0.001**	[0.11, 0.47]
RMSE	-0.13	0.80	-0.16	0.869	[-1.71, 1.44]
Spectral centroid	0.00	0.00	4.62	**<0.001**	[0.00, 0.00]
Brightness	-0.94	0.24	-3.88	**<0.001**	[-1.42, -0.47]
Dissonance	0.00	0.00	-2.97	**0.003**	[0.00, 0.00]
Mode	0.28	0.11	2.59	**.010**	[0.07, 0.50]
Event density	-0.03	0.01	-2.77	**.006**	[-0.06, -0.01]
Mean pitch	0.00	0.00	-0.17	0.864	[0.00, 0.00]
ZCR	0.00	0.00	-1.82	0.070	[0.00, 0.00

Table 7. Regression statistics for valence model.

and brightness), and dynamic (RMSE) musical features. All significant music predictors had small effect sizes except for tempo.

Regression model: Valence
This model used all 749 observations derived from the relevant music excerpts included in this study. Residuals for the overall model were normally distributed according to the Shapiro-Wilk test ($p > .05$). Furthermore, the assumption of linearity was met after inspection of the plot of the residual/fitted values. Slight clustering towards the centre of the residuals was observed; however, White's test indicated the assumption of homoscedasticity was not violated ($p = .109$).

The general model determining if musical attributes are associated with perceptions of valence in music was significant $F(10, 738) = 11.42$, $p < .001$, explaining 13.40% of the variance in perceived impressions of valence. Table 7 shows three significant positive relationships: valence and pulse clarity ($\eta_p^2 = 0.013$), valence and spectral centroid ($\eta_p^2 = 0.028$) and valence and mode ($\eta_p^2 = 0.010$). Therefore, positive valence was associated with a more discernible pulse in the music, a higher spectral centroid and a more positive value for modality (indicating that the music excerpt was more likely to be in a major key), holding all other variables constant.

The model revealed three significant negative relationships: valence and brightness ($\eta_p^2 = 0.020$), valence and dissonance ($\eta_p^2 = 0.012$) and valence and event density ($\eta_p^2 = 0.010$) (see Table 7). These relationships indicate that a one-unit increase in either brightness, event density, or dissonance was associated with decreased pleasantness. That is, less pleasant valence ratings were related to increased high-frequency energy, higher levels of dissonance, and a higher number of musical events in a segment. Overall, perceptions of valence in music were influenced by multiple broad musicological concepts such as rhythm (event density and pulse clarity), harmony (mode and dissonance), and timbre (brightness and spectral centroid).

Models of musical attributes and perceived social emotions
Regression model: Dominance
The residuals were normally distributed according to the Shapiro-Wilk test ($p > 0.05$), and the assumption of linearity was met after inspection of the residuals against the plot of the fitted values. There was very slight clustering towards the centre of the residual plot. However, White's test indicated that the assumption of homoscedasticity was not violated ($p = .142$).

The model for musical attributes influencing the perception of dominance in music was significant, $F(10, 738) = 170.35$, $p < .001$, explaining 69.77% of the variance in perceptions of dominance in music. Three significant positive associations were observed: dominance and tempo ($\eta_p^2 = 0.484$), dominance and brightness ($\eta_p^2 = 0.018$) and dominance and dissonance ($\eta_p^2 = 0.026$) (see Table 8). That is, higher ratings of dominance

5

Variable	Coef.	SE	t	p	95% CI
Tempo	0.65	0.02	26.31	**< 0.001**	[0.60, 0.70]
Pulse clarity	-0.24	0.07	-3.55	**< 0.001**	[-0.38, -0.11]
RMSE	-1.06	0.61	-1.74	0.083	[-2.26, 0.14]
Spectral centroid	0.00	0.00	-3.68	**< 0.001**	[0.00, 0.00]
Brightness	0.67	0.18	3.64	**< 0.001**	[0.31, 1.03]
Dissonance	0.00	0.00	4.44	**< 0.001**	[0.00, 0.00]
Mode	0.00	0.08	0.05	0.964	[-0.16, 0.17]
Event density	-0.01	0.01	-0.95	0.343	[0.00, 0.00]
Mean pitch	0.00	0.00	1.68	0.094	[0.00, 0.00]
ZCR	0.00	0.00	−0.033	0.743	[0.00, 0.00

Table 8. Regression statistics for dominance model.

Variable	Coef.	SE	t	p	95% CI
Tempo	0.33	0.03	12.01	**< 0.001**	[0.27, 0.38]
Pulse clarity	0.34	0.08	4.52	**< 0.001**	[0.19, 0.49]
RMSE	1.24	0.67	1.85	0.064	[-0.07, 2.56]
Spectral centroid	0.00	0.00	3.54	**< 0.001**	[0.00, 0.00]
Brightness	-0.37	0.20	-1.82	0.069	[-0.77, 0.03]
Dissonance	0.00	0.00	-3.03	**0.003**	[0.00, 0.00]
Mode	0.28	0.09	3.08	**0.002**	[0.10, 0.46]
Event density	-0.04	0.01	-3.58	**< 0.001**	[-0.06, 0.02]
Mean pitch	0.00	0.00	-0.11	0.910	[0.00, 0.00]
ZCR	0.00	0.00	-2.11	0.035	[0.00, 0.00]

Table 9. Regression statistics for affiliation model.

were assigned to music excerpts that were faster in tempo, had greater levels of high-frequency energy, and had higher levels of dissonance.

The model also revealed two significant negative relationships between perceptions of dominance and pulse clarity ($\eta_p^2 = 0.017$) and perceptions of dominance and spectral centroid ($\eta_p^2 = 0.018$) (see Table 8). Therefore, when the pulse of the music was more discernible or the spectral centroid value was higher, music excerpts were likely to be perceived as less dominant. Overall, the perception of dominance was influenced by rhythmic (tempo and pulse clarity), harmonic (dissonance) and timbral (brightness and spectral centroid) features of music. While these effects were found to be highly significant, the effect sizes were small, except for tempo.

Regression model: Affiliation
This model included all 749 music excerpt observations. The assumption of normality of residuals was not met according to the Shapiro-Wilk test ($p < .05$). After inspection of a residual against fitted values plot, the assumption of linearity was met; however, the assumption of constant variance may be violated due to moderate clustering towards the middle of the plot. This was confirmed by White's test, indicating that the assumption of homoscedasticity was violated ($p = .032$). Despite this potential caveat, we continued with a multiple regression model due to the large sample size and the robust nature of multiple linear regression.

The overall model explaining the relationship between musical attributes and perceived affiliation in music was significant, $F(10, 738) = 47.53$, $p < .001$, explaining 39.18% of perceived affiliation in music. As seen in Table 9, tempo ($\eta_p^2 = 0.163$), pulse clarity ($\eta_p^2 = 0.027$), spectral centroid ($\eta_p^2 = 0.017$), and mode ($\eta_p^2 = 0.013$) all had a significant positive relationship with the perception of affiliation. Thus, a one-unit increase in any of these predictors while other variables are held constant results in more social perceptions of affiliation. Music excerpts that were higher in tempo, had a more discernible pulse, had a higher spectral centroid value, or a higher likelihood of being in a major key were more likely to be interpreted as prosocial.

Dissonance ($\eta_p^2 = 0.012$) and event density ($\eta_p^2 = 0.017$) had a significant negative association with the perception of affiliation (see Table 9). When controlling for other variables, a one-unit increase in either of these predictors results in decreased perceptions of social affiliation in music. In turn, music excerpts with greater amounts of dissonance or several musical events per segment were more likely to be perceived as antisocial. However, according to the coefficients, confidence intervals, and effect size, the effect of dissonance and spectral centroid were small. Effect sizes for all other significant predictors, excluding tempo, were small. Overall, the model suggests that the perception of affiliation in music is related to rhythmic (tempo, event density and pulse clarity), harmonic (dissonance and mode), and timbral (spectral centroid) features of the music.

Discussion

This investigation considered whether social emotions (dominance and affiliation) and emotional dimensions of EA, TA and valence are communicated by brief music excerpts across a large corpus of genres and whether musical attributes of excerpts (e.g. rhythm, timbre, harmony) predict the emotions perceived. Previous research has confirmed that individuals perceive dominant or affiliative social interactions between performing musicians[38] but not whether social emotions are perceived by listeners in music across genres.

The concepts of dominance and affiliation are central to social interactions and have been extensively studied in psychological literature. However, they have not been examined in relation to music experience. Dominance relates to feelings of control, assertiveness, and influence, whereas affiliation refers to feelings of connection. That music can signal these social constructs is a novel finding.

Figure 1 provides a framework for interpreting the findings. The figure displays the music features identified in the corpus of music samples and their association with perceived social emotions and the emotional dimensions of EA, TA and valence. As depicted in the figure and supported in regression analyses, multiple musical attributes converge to predict specific emotional qualities, with different sets of attributes combining to predict different emotional qualities. These different interconnections, along with evidence for the communication of these emotions, will next be discussed with respect to existing models of music and emotion.

Multiple regression models revealed that ratings of dominance and affiliation were significantly predicted by music variables, with a higher percentage of variance explained for dominance (69.77%) than affiliation (39.18%). Therefore, the hypothesis that musical attributes can be grouped and classified according to their impact on ratings of social emotions was supported.

These findings complement the suggestion by Aucouturier and Canonne[38] that music is perceived as the sonic trace of social relations among musicians, and that the acoustic consequences of these relationships help listeners infer social intentions like dominance and affiliation. Their experimental approach involved music improvisers communicating non-musical social intentions, and they found that these intentions were recognised based on dynamic interactions and specific acoustic cues. In contrast, our study focused on the capacity of static (average) musical attributes to predict the perception of social emotions and revealed that while both dominance
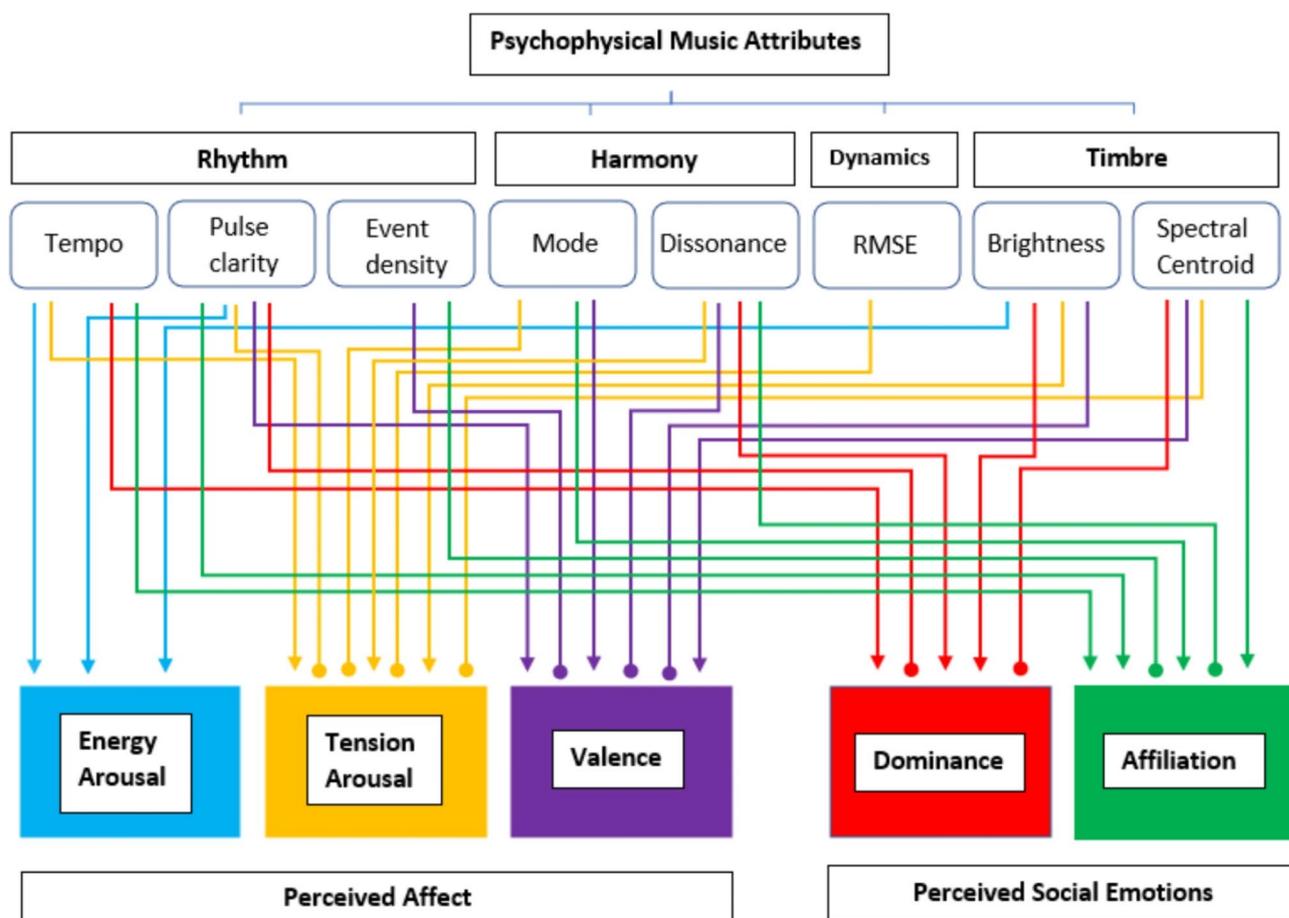

Fig. 1. Relationships between music features and rating scales. *Note:* All connections represent significant psychoacoustic predictors in each model. Musical attributes explained the following percentages of variance: Energy Arousal (82.28%), Tension Arousal (58.02%), Valence (13.40%), Dominance (69.77%), and Affiliation (39.18%). Lines ending with a triangle represent positive relationships; lines ending with a circle represent negative relationships. RMSE = root mean square energy.

and affiliation can be predicted by such features, the degree of prediction varied. Specifically, musical attributes explained 30.59% more variance in the dominance model than in the affiliation model.

Further research is needed to elucidate sources of variance in social emotions that are not explained by musical attributes, especially in the case of affiliation. To the extent that music is sometimes employed as a surrogate friend[56], it is likely that feelings of affiliation are triggered by enjoyment of familiar music. Bivariate correlations revealed that perceived dominance significantly correlated to all music predictors (except for modality), before subsequent music feature reduction via PCA. Moreover, perceived affiliation was correlated with all musical attributes except for dissonance before subsequent music feature reduction.

Many structural features of music were associated with the social dimensions of affiliation and dominance. The magnitude of perceived dominance in music was related to rhythmic (tempo and pulse clarity), harmonic (dissonance), and timbral (brightness and spectral centroid) music features. The magnitude of perceived affiliation in music was related to somewhat different musical predictors but included rhythmic (tempo, pulse clarity, and event density), harmonic (dissonance and modality), and timbral (spectral centroid) features. Given violations of the assumption of normality and homoscedasticity in our analysis, further research on such models is warranted.

Each dimensional model of perceived affect (EA, TA, and valence) was highly significant ($p < .001$), and dimensional measures of affect were significantly correlated with each other. This observation is consistent with previous research that ratings of valence and tension are negatively correlated ($r = -.70$), whereas ratings of energy and tension are positively correlated ($r = .57$)[12,39]. The explained variance for the EA and TA models was larger than the explained variance for the valence model. This finding is consistent with previous evidence that musical attributes are more reliably associated with arousal ratings than valence ratings[40,41].

The difficulty in explaining valence ratings may be partly explained by the large diversity of genres included. Different genres likely evoked distinct emotional responses and personal associations, which may have masked a consistent pattern in how musical attributes relate to valence. Factors such as familiarity with pieces or genres, personal associations, and visual imagery likely contributed to judgments, reducing the predictive power of musical attributes across genres[42].

Figure 1 illustrates significant associations. Tempo was the strongest predictor of perceived EA and TA. The effect size was large for both EA and TA, consistent with previous research on the association between tempo and arousal[34,43]. Moreover, music excerpts with a more discernible beat (higher pulse clarity) were associated with increased energy but decreased tension. Event density was not related to either dimension.

Musical attributes that constitute harmony (mode and dissonance) were significantly associated with perceptions of TA. Music with high levels of dissonance or in a minor key yielded high tension ratings by listeners (note that the coefficient of dissonance was small and rounded to 0.00. MIRtoolbox extracts values for dissonance that range from 3 to 1390, so coefficients for predictions of ratings on a 7-point scale are small). Increased RMSE (loudness) was significantly associated with decreased perceptions of tension in music. However, the zero-crossing rate was not a significant predictor of arousal.

As predicted, higher magnitudes of perceived energy in music were significantly predicted by increased brightness. This supports previous findings of the positive relationship between brightness and arousal[44]. Furthermore, tension was also significantly predicted by spectral centroid, revealing that the higher the spectral distribution centroid of a sample, the less likely it will express tension (note that spectral centroid displays similarly small coefficient values to those for dissonance (rounded to 0.00), as spectral centroid values range from 405 to 6076).

Two of the three musical attributes related to rhythm were significant predictors of perceived pleasantness in music. Tempo did not predict perceived pleasantness in music. Although previous research suggests that tempo predicts judgments of happiness and sadness[25,45,46] (but see[43]), its lack of association with ratings of pleasantness may partly reflect the fact that both happy and sad music can be pleasing[2,11]. That is, the (dimensional) rating scale for pleasantness used in this study may diverge from ratings based on categorical models of valence (e.g. happy, joyful, or sad). Pulse clarity and event density were important for predicting perceived pleasantness in music, albeit with small effect sizes. The directionality of these associations suggests that music with a more discernible beat (pulse clarity) is likely to increase perceived pleasantness for the listener. In contrast, music that contains more notes per musical moment (event density) is likely to decrease perceived pleasantness for the listener.

Both structural features related to harmony (mode and dissonance) predicted perceived pleasantness in music. These relationships reveal that musical excerpts in a major mode and with less harsh sounds are more likely to be perceived as pleasant[21,22]. In contrast to previous findings[21], the mean pitch was not a significant predictor of perceived pleasantness in music. One interpretation is that the MIRtoolbox could not extract reliable measures of mean pitch from such brief (5-second) samples, suggesting the need to revisit this predictor based on longer music samples.

RMSE (loudness) and zero-crossing rate were not significant predictors of perceived pleasantness in music, suggesting that variations in loudness and transient information may not play a crucial role in shaping listeners' perceptions of pleasantness. This finding can be attributed to the standardised sound level of most MP3 files used for audio feature extraction, which may have masked the potential impact of loudness variation. Most excerpts returned RMSE values between 0.05 and 0.15 (range 0.03 and 0.25), resulting in insufficient variability in loudness to detect an association with perceived pleasantness.

Musical attributes associated with timbre, such as brightness and spectral centroid, emerged as significant predictors of pleasantness. Specifically, the higher an audio sample's spectral distribution centroid, the more likely it was perceived as pleasant. Moreover, increases in high-frequency energy in the audio sample was associated with reduced pleasantness, implying that brightness or 'sharpness' in music may be perceived as unpleasant. These findings illustrate the relevance of timbral qualities for the judgments of pleasantness in music.

Given the high correlation between dominance and energy, a question arises whether listeners equated the two concepts. This possibility is unlikely, however, as steps were taken to ensure that participants fully understood the distinction between dominance and arousal. Arousal is a general emotional state characterized by physiological activation and heightened alertness and can be experienced across a range of contexts, whether social or non-social. It reflects the intensity of emotional experience but is not tied to social dynamics. In contrast, dominance is a social emotion that describes feelings of power or control within social situations. Unlike arousal, which can occur in isolation or in response to specific stimuli, dominance is relational in nature, emerging in social interactions where hierarchical positioning or authority is at play. Our instructions to respondents clarified that dominance is a social emotion, and not merely a heightened state of arousal. This distinction was crucial to capture the unique emotional responses related to dominance, independent of arousal.

The models incorporating structural features of music predicted both social emotions and the emotional dimensions of EA, TA, and valence. Moreover, musical attributes were strong predictors of energy, tension, and dominance, but comparatively weak predictors of valence and affiliation. These findings raise the question of whether affiliation, as a social emotion, is spontaneously perceived in music without explicit prompting or social conditions. Affiliation, which involves feelings of connection, warmth, and social bonding, may not be encoded exclusively within specific musical elements, especially in the absence of contextual or lyrical cues that indicate social interaction or togetherness. Conversely, feelings of affiliation may be primed by approaches to listening that treat music as a 'surrogate friend'[56]. Future research should aim to identify the musical features, contexts, and priming conditions that enhance or diminish the perception of affiliation in music.

Nonetheless, it is notable that affiliation was better explained than valence, suggesting that the model for affiliation has a stronger predictive power than that for valence. It may be valuable in future research to examine how judgments of affiliation and dominance relate to core affect dimensions, such as pleasure-displeasure and activation-deactivation. Understanding these relationships could enhance the theoretical framework of emotion perception in music.

A limitation of this study – and that of Cowen et al.'s study from which stimuli were drawn[1]– was the brief duration of each stimulus. Presenting brief music samples permitted many excerpts to be presented while avoiding potential participant fatigue[23,24], but it is uncertain how much emotional information is contained within such a short sample. Research has confirmed that listeners can detect an emotional character in brief musical excerpts[1,4,57], but complex emotions such as nostalgia and awe may take longer than 5 s to convey[42]. The brief duration of stimuli likely restricted the kinds of emotions communicated.

A second limitation is that participants were asked to rate all measures of affect and social emotions sequentially. This procedure required that each musical stimulus be continually and retrospectively judged in terms of energy, tension, valence, dominance, and affiliation, which may have been difficult given the small amount of musical information provided. Interestingly, listeners did not appear to have difficulty with the rating scales, and examination of the data indicated that all rating scales were systematically related to musical attributes. These findings suggest that the procedure did not interfere with the quality of the results.

Our recruitment strategy also limited the collection of detailed demographic information, except that all participants were older than 18 and from the United States of America. It is uncertain how population parameters such as age, gender, or previous musical experience relate to emotion perception. Such information, though not the purpose of the investigation, could provide additional insight into the connection between music and emotion and remains an important issue for future research.

The finding that social emotions are perceived in music may help to explain the strong motivation to engage with music. The associations identified between social emotions and structural features of music, in turn, have implications for music composition, film scoring and musical theatre. Evidence that dominance and affiliation are perceived in music raises a question of what mechanisms are involved in this effect and how these social emotions are conveyed over extended periods. Time-series analysis could be explored to investigate the dynamic properties of dominance and affiliation, perhaps revealing that the social connotations of music wax and wane in intensity as music unfolds over time.

To conclude, this investigation lays a foundation for future research on social emotions in music, with implications for social well-being and intercultural understanding. Participants readily perceived the emotional connotations of dominance and affiliation within short music excerpts, indicating that music can encode social emotions. Such findings are compatible with theoretical discussions of music that emphasize its social functions[2,11,38,56,58]. Judgments of dominance were strongly predicted by musical attributes, whereas judgments of affiliation were weakly associated with such attributes. The findings also highlight the important role of timbral qualities in shaping perceptions of pleasantness in music. Overall, the research provides new evidence that expands our theoretical understanding of how social emotions are perceived in music.

Method
Participants
Participants were recruited from the crowdsourcing platform Amazon Mechanical Turk (MTurk), with responses from 1563 participants recorded. Participants were based in the United States, were English-speaking, and were at least 18 years old. All participants provided consent after being informed of the aims of the survey.

Sampling procedure
Data collection occurred in July 2021. Participants completed the survey for compensation of $0.30 USD. To improve quality, MTurk workers were prequalified as residing in the USA, previously completing a minimum of 50 MTurk tasks and achieving a prior minimum MTurk approval rating of 90%.

Materials and measures

Stimulus set

Seven hundred and fifty five-second music instrumental excerpts (no spoken words or lyrics included) were gathered and used as test stimuli from an existing database[1]. The corpus of music comprised 16 genres (music metadata): Alternative ($n=38$), Ambient ($n=95$), Classical ($n=94$), Country ($n=21$), EDM ($n=82$), Electronic ($n=81$), Folk ($n=24$), Heavy Metal ($n=37$), Hip-hop ($n=15$), Jazz ($n=31$), Latin ($n=7$), Pop ($n=64$), Rock ($n=82$), R&B ($n=24$), Reggae ($n=3$), and World Music ($n=52$). Note that 'World Music' includes a diverse range of musical traditions from across the globe, encompassing both non-Western and less mainstream Western genres such as traditional music of the Scottish Highlands.

Randomisation of stimulus subsets presented to participants

From the 750 music excerpts, 50 survey subsets were generated, each containing 15 music excerpts. Each participant was randomly assigned one of the subsets of 15 music excerpts in Qualtrics and responded to all excerpts in the subset to qualify for compensation. Randomisation was constrained in two ways. First, it was evenly distributed across survey subsets from the beginning of data collection, which was accomplished by requiring at least one participant to be allocated to every survey subset and their responses submitted before allocating a second participant to a particular subset. This process was iteratively repeated until all participants were allocated and responses were received. Second, no genre was represented more than three times in any one survey subset, ensuring that participants heard a range of musical genres in their subset.

Measures

Acoustic measures

MIRtoolbox. Acoustic analyses were performed using the music information retrieval (MIR) toolbox[47] (see Fig. 2).

   MIR root mean square energy (RMSE). A measure of amplitude (loudness), taken as the global energy of the audio sample, was estimated by taking the square root of the average squared amplitude over a period of time[47]. Scores range from 0 to 1, with higher scores indicative of a louder signal in the audio sample.

   MIR brightness. Brightness measures the energy in an audio sample above a frequency[47,51]. Scores range from 0 to 1, with higher scores signalling greater high-frequency energy in the sample.

   Roughness. Roughness is a measure of global sensory dissonance, corresponding to the "pulsing" phenomenon experienced when multiple sounds slightly deviate in frequency are heard simultaneously[47]. Sensory dissonance is estimated by taking the average of all pairs of peaks in an audio sample frequency spectrum. Scores range between 3 and 1390, with higher scores indicating increased sensory dissonance and, as a result, more harsher sounds in the audio sample.

   Rolloff. This measures how much energy in an audio sample is below an energy threshold[47]. MIRtoolbox uses a percentage cut-off of 0.85[49]. Higher values on the rolloff measure indicate that more energy is confined below the threshold.
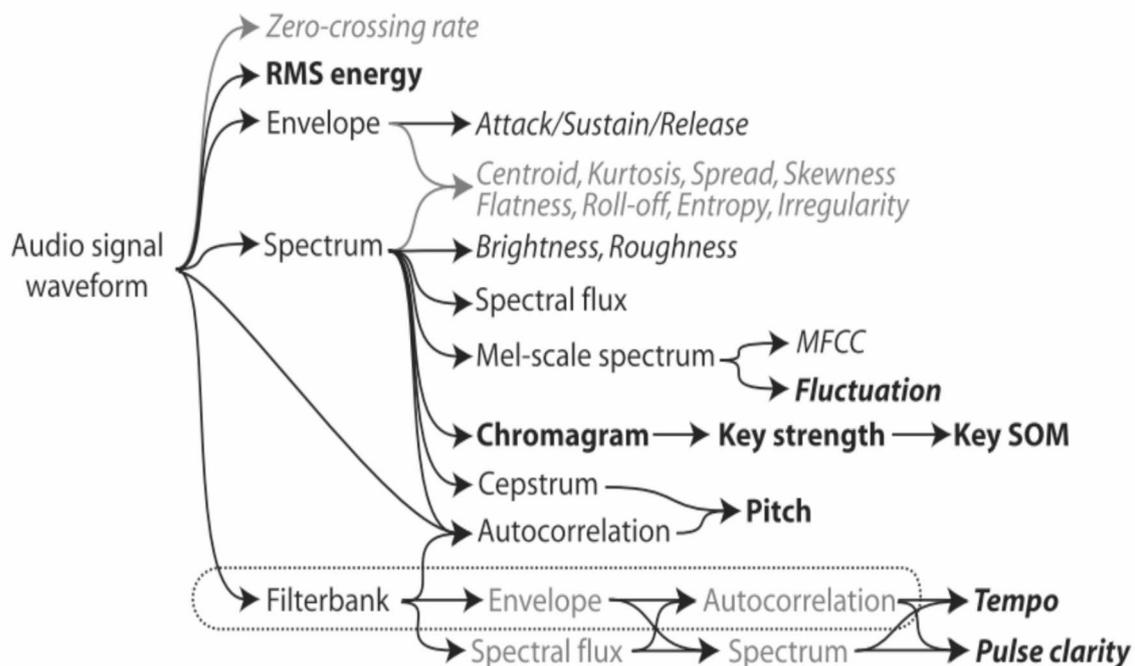

Fig. 2. MIRtoolbox computational methods for audio feature extraction (Lartillot et al., 2008).

MIR zerocross. The zero-crossing rate measures the amount of noisiness in an audio sample[49]. It is estimated by counting the times the signal crosses the X-axis (zero amplitude)[47,50]. Higher values indicate a greater zero-crossing rate and a noisier audio sample.

MIR mode. A measure of whether the audio sample is either major or minor in modality[47]. Values closer to 1 indicate audio samples are more likely to be in a major key − 1 suggests a minor key. A value close to 0 indicates that the modality of the audio sample is ambiguous.

MIR spectral centroid. A measure of an audio sample's spectral distribution centroid[47]. Spectral centroid is related to brightness, with higher values indicating lower brightness (e.g., an oboe's spectral centroid is higher than a French horn's)[51].

MIR pitch. A measure of mean pitch using autocorrelation to estimate the average frequency of all pairs of peaks in the audio sample[47,52]. Higher values indicate a greater presence of sounds.

MIR pulse clarity. Pulse clarity is the musical measure of how easily the listener can identify the underlying beat or metrical pulsation in a given piece of music[47]. Scores range from 0 to 1, with higher scores indicating that the listener easily discerns the underlying beat or metrical pulsation.

MIR event density. Event density measures the number of events (peaks) occurring in the audio sample and approximates the mean frequency of these events per second[47,52]. Higher values indicate an increase of notes in each musical moment.

MIR spread; skewness; kurtosis; flatness; entropy (spectral properties). These five measures all relate to the spectral dispersion of an audio sample. MIR spread is the standard deviation of the spectral distribution. MIR flatness calculates the flatness of data, indicating if the spectral distribution is smooth or peaky. Finally, MIR entropy represents the respective Shannon entropy of the spectrum[53], with higher values signalling more uncertainty and peakiness.

Response measures
The circumplex model of affect classifies emotion according to the dimensions of arousal and valence[17], but arousal was further subdivided into energy arousal (EA) and tension arousal (TA)[12,54]. Two *social* emotions, dominance and affiliation, were also included[28]. Thus, the study examined three emotional dimensions of EA, TA, and valence, and two social emotions of dominance and affiliation.

Participants rated the emotional quality of the music. Each bipolar rating scale ranged from one to seven, with ratings five and above indicating high levels of emotion and ratings of three and below indicating low levels of emotion. *Energy arousal* ranged from tired to energetic; *tension arousal* ranged from relaxed to tense; *valence* ranged from unpleasant to pleasant; *dominance* ranged from submissive to dominant; and *affiliation* ranged from antisocial to highly social. As a rating of four represented a neutral option for all scales, it was possible to assess whether mean ratings for each dimension (e.g., affiliation, dominance, valence) differed significantly from a neutral option.

Modified Self-Assessment Manikin (SAM). The SAM measures an individual's pleasure, arousal, and dominance by displaying visual depictions of the different levels of each measure using schematic manikins comparable to emojis[55]. The current study adopted this strategy to explain and measure the dimensions of dominance and affiliation, emphasizing their social meaning. Dominance and affiliation were conceptualised by two scales containing five graphical figures, each representing a 7-point response scale for each dimension. Dominance ranged from − 3 (illustrated by a small vulnerable figure in a large open space) to 3 (depicted by a large figure dominating the surrounding space). Affiliation ranged from − 3 (illustrated by an isolated figure from a group) to 3 (depicted by the same figure now in the middle of the group).

Tempo. Tempo ratings on a 7-point scale (7 = fast) were included to assess participants' perceptions of how fast the music excerpts were. This measure was included to corroborate the MIRtoolbox extraction of tempo, given the samples were so short.

Procedure
The survey was published on MTurk and redirected participants to a Qualtrics survey. Before completing the survey, participants read a project description and provided informed consent. They were informed that financial compensation would only be provided if they responded to all 15 audio samples in the survey. This research was approved by the Macquarie University Human Research Ethics Committee (HREC), and all methods were performed in accordance with their ethical guidelines and regulations, as well as the publishing policies of *Scientific Reports*.

Participants were instructed to remember that the study investigated the emotions we *perceive* in music, not emotions that might be *induced* by music. This distinction is important as music can have emotional connotations (e.g., sadness). Moreover, the nature of each bipolar response scale was explained with responses on the far left and far right indicating an intense perception of the dimension in question (e.g., unpleasant, pleasant). Afterwards, instructions for judging the dimensions of *affiliation* and *dominance* were provided. Participants were instructed to imagine the music as a film score that accompanied a character in the film, and to judge what that score suggested about the characteristics of the character. For affiliation, listeners judged whether the music suggested a social or antisocial character. For dominance, listeners judged whether the music suggested a dominant or submissive character. The modified SAM was also presented alongside these instructions to aid the participants' understanding of the concepts of *affiliation* and *dominance* in music. Participants were informed they could listen to each stimulus twice. Participants were allowed one hour to complete the 15 items in the survey.

After listening to the audio sample, participants were instructed to make six ratings: (1) how tired or energetic the audio sample was; (2) how relaxed or tense the audio sample was; (3) how pleasant or unpleasant the audio sample was (4) how submissive or dominant the audio sample was; (5) how antisocial or social the audio sample

was; and (6) how slow or fast the audio sample was. Participants had to answer all six response scales before proceeding to the next page. This process was repeated for each of the 15 music stimuli.

Data screening

750 music excerpts were included in this study, with six response scales associated with each excerpt. One music excerpt was excluded from the study as its mean pitch value could not be calculated. Thus, 749 music excerpts were included in the statistical analysis.

1563 participants had their responses recorded. Participants were excluded if they did not give consent ($n = 3$), were judged to have been bots ($n = 15$) based on reCAPTCHA score, did not read instructions ($n = 11$), did not fully respond to the survey ($n = 7$), or responded with only one number ($n = 14$). After these 50 participants had been removed from the dataset, 1513 participants remained for subsequent statistical analysis.

After participant exclusion, randomization into survey subsets resulted in 27 to 32 participants allocated into each survey subset. Therefore, each response scale associated with a music excerpt was answered a minimum of 27 times and a maximum of 32 times. As responses to the 749 music excerpts were distributed across different sets of participants, music excerpts, rather than participants, were treated as the random variable in statistical analyses. In other words, the participants' responses generated a mean rating scale value for perceived EA, TA, valence, dominance, affiliation, and tempo for each music excerpt. This allows for each excerpt to contain values for both music features (obtained through MIRtoolbox) and perceived ratings of EA, TA, valence, dominance, and affiliation for subsequent statistical analysis.

Data availability

Data are available from either W.F.T. or K.N.O., and held at Macquarie University and Bond University.

References

1. Cowen, A. S., Fang, X., Sauter, D. & Keltner, D. What music makes us feel: at least 13 dimensions organize subjective experiences associated with music across different cultures. *Proc. Natl. Acad. Sci.* **117** (4), 1924–1934. https://doi.org/10.1073/pnas.1910704117 (2020).
2. Juslin, P. N. *Musical Emotions Explained: Unlocking the Secrets of Musical Affect* (Oxford University Press, 2019).
3. Juslin, P. N. & Laukka, P. Expression, perception, and induction of musical emotions: a review and a questionnaire study of everyday listening. *J. New. Music Res.* **33** (3), 217–238. https://doi.org/10.1080/0929821042000317813 (2004).
4. Day, R. A. & Thompson, W. F. Measuring the onset of experiences of emotion and imagery in response to music. *Psychomusicology: Music Mind Brain.* **29** (2–3), 75–89. https://doi.org/10.1037/pmu0000220 (2019).
5. Gabrielsson, A. & Juslin, P. N. Emotional expression in music. In (eds Davidson, R. J., Scherer, K. R. & Goldsmith, H. H.) Handbook of Affective Sciences (503–534). Oxford University Press. (2003).
6. Juslin, P. N. What does music express? Basic emotions and beyond. *Front. Psychol.* **4**, 1–14. https://doi.org/10.3389/fpsyg.2013.00596 (2013).
7. Sznycer, D., Sell, A. & Lieberman, D. Forms and functions of the Social emotions. *Curr. Dir. Psychol. Sci.* **30** (4), 292–299. https://doi.org/10.1177/09637214211007451 (2021).
8. Tangney, J. P., Stuewig, J. & Mashek, D. J. Moral emotions and moral behavior. *Ann. Rev. Psychol.* **58**, 345–372. https://doi.org/10.1146/annurev.psych.56.091103.070145 (2007). PMID: 16953797; PMCID: PMC3083636.
9. Camacho-Morles, J. et al. Activity achievement emotions and academic performance: a meta-analysis. *Educational Psychol. Rev.* 1–45. https://doi.org/10.1007/s10648-020-09585-3 (2021).
10. Vogl, E., Pekrun, R. & Loderer, K. Epistemic emotions and metacognitive feelings. In D. Moraitou & P. Metallidou (Eds.), Trends and prospects in metacognition research across the life span: A tribute to Anastasia Efklides (pp. 41–58). Springer Nature Switzerland AG. (2021). https://doi.org/10.1007/978-3-030-51673-4_3
11. Thompson, W. F., Bullot, N. J. & Margulis, L. H. The psychological basis of music appreciation: structure, self, source. *Psychol. Rev.* **130** (1), 260–284. https://doi.org/10.1037/rev0000364 (2023).
12. Eerola, T. & Vuoskoski, J. K. A comparison of the discrete and dimensional models of emotion in music. *Psychol. Music.* **39** (1), 18–49. https://doi.org/10.1177/0305735610362821 (2011).
13. Yang, Y. H. & Chen, H. H. Machine recognition of music emotion: a review. *ACM Trans. Intell. Syst. Technol.* **3** (3). https://doi.org/10.1145/2168752.2168754 (2012).
14. Olsen, K. N., Terry, J. & Thompson, W. F. Psychosocial risks and benefits of exposure to heavy metal music with aggressive themes: current theory and evidence. *Curr. Psychol.* **42**, 21133–21150. https://doi.org/10.1007/s12144-022-03108-9 (2023).
15. Thompson, W. F., Geeves, A. M. & Olsen, K. N. Who enjoys listening to violent music and why? *Psychol. Popular Media Cult.* **8** (3), 218–232. https://doi.org/10.1037/ppm0000184 (2019).
16. Cohrdes, C., Wrzus, C., Wald-Fuhrmann, M. & Riediger, M. The sound of affect: Age differences in perceiving valence and arousal in music and their relation to music characteristics and momentary mood. *Musicae Sci.* **24** (1), 21–43. https://doi.org/10.1177/1029864918765613 (2018).
17. Russell, J. A. A circumplex model of affect. *J. Personal. Soc. Psychol.* **39** (6), 1161–1178. https://doi.org/10.1037/h0077714 (1980).
18. Russell, J. A. & Barrett, L. F. Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *J. Personal. Soc. Psychol.* **76** (5), 805–819. https://doi.org/10.1037/0022-3514.76.5.805 (1999).
19. Ilie, G. & Thompson, W. F. A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Percept.* **23**, 319–329 (2006).
20. Thayer, R. E. *The Biopsychology of mood and Arousal* (Oxford University Press, 1989).
21. Juslin, P. N. & Lindström, E. Musical expression of emotions: modelling listeners' judgements of composed and performed features. *Music Anal.* **29** (1–3), 334–364. https://doi.org/10.1111/j.1468-2249.2011.00323.x (2010).
22. Carr, N. R., Olsen, K. N. & Thompson, W. F. The perceptual and emotional consequences of articulation in music. *Music Percept.* **40**, 202–219. https://doi.org/10.1525/mp.2023.40.3.202 (2023).
22. Laurier, C., Lartillot, O., Eerola, T. & Toiviaine, P. Exploring relationships between audio features and emotion in music. Conference Abstract: Tuning the Brain for Music, 260–264. (2009). https://doi.org/10.3389/conf.neuro.09.2009.02.033
23. Eerola, T., Friberg, A. & Bresin, R. Emotional expression in music: contribution, linearity, and additivity of primary musical cues. *Front. Psychol.* **4**, 1–12. https://doi.org/10.3389/fpsyg.2013.00487 (2013).

23. Skowronek, J., McKinney, M. F. & Van De Par, S. Ground truth for automatic music mood classification. ISMIR 2006 7th International Conference on Music Information Retrieval, 395–396. (2006).

23. Yang, Y. H., Lin, Y. C., Su, Y. F. & Chen, H. H. A regression approach to music emotion recognition. *IEEE Trans. Audio Speech Lang. Process.* **16** (2), 448–457. https://doi.org/10.1109/TASL.2007.911513 (2008).

24. Panda, R., Rocha, B. & Paiva, R. P. Music emotion recognition with standard and melodic audio features. *Appl. Artif. Intell.* **29** (4), 313–334. https://doi.org/10.1080/08839514.2015.1016389 (2015).

25. Bresin, R. & Friberg, A. Emotion rendering in music: Range and characteristic values of seven musical variables. *Cortex.* **47** (9), 1068–1081. https://doi.org/10.1016/j.cortex.2011.05.009 (2011).

26. Hareli, S., David, S. & Hess, U. The role of emotion transition for the perception of social dominance and affiliation. *Cognition Emot.* **30** (7), 1260–1270. https://doi.org/10.1080/02699931.2015.1056107 (2016).

27. Hess, U., Blairy, S. & Kleck, R. E. The influence of expression intensity, gender, and ethnicity on judgments of dominance and affiliation. *J. Nonverbal Behav.* **24**, 265–283. https://doi.org/10.1023/A:1006623213355 (2000).

28. Mobbs, A. E. D. An atlas of personality, emotion and behaviour. *PLoS ONE.* **15** (1), 1–25. https://doi.org/10.1371/journal.pone.0227877 (2020).

29. van Kleef, G. A. & Côté, S. The Social effects of emotions. *Ann. Rev. Psychol.* **73**, 629–658. https://doi.org/10.1146/annurev-psych-020821-010855 (2022).

30. Quirin, M. et al. Neural correlates of social motivation: an fMRI study on power versus affiliation. *Int. J. Psychophysiol.* **88** (3), 289–295. https://doi.org/10.1016/j.ijpsycho.2012.07.003 (2013).

31. Mehrabian, A. Pleasure-arousal-dominance: a general framework for describing and measuring individual differences in temperament. *Curr. Psychol.* **14** (4), 261–292. https://doi.org/10.1007/bf02686918 (1996).

32. Mehrabian, A. & Russell, J. A. Evidence for a three-factor theory of emotions. *J. Res. Pers.* **11** (3), 273–294. https://doi.org/10.1016/0092-6566(77)90037-X (1977).

33. Russell, J. A. Evidence of convergent validity on the dimensions of affect. *J. Personal. Soc. Psychol.* **36** (10), 1152–1168. https://doi.org/10.1037/0022-3514.36.10.1152 (1978).

34. Brinker, B., Den, Dinther, R., Van & Skowronek, J. Expressed music mood classification compared with valence and arousal ratings. *Eurasip J. Audio Speech Music Process.* **24**, 1–14. https://doi.org/10.1186/1687-4722-2012-24 (2012).

35. Gingras, B., Marin, M. M. & Fitch, W. T. Beyond intensity: spectral features effectively predict music-induced subjective arousal. *Q. J. Experimental Psychol.* **67** (7), 1428–1446. https://doi.org/10.1080/17470218.2013.863954 (2014).

36. Grekow, J. Audio features dedicated to the detection and tracking of arousal and valence in musical compositions. *J. Inform. Telecommunication.* **2** (3), 322–333. https://doi.org/10.1080/24751839.2018.1463749 (2018).

37. Thompson, W. F., Schellenberg, E. G. & Husain, G. Arousal, mood, and the Mozart Effect. *Psychol. Sci.* **12** (3), 248–251 (2001).

38. Aucouturier, J. J. & Canonne, C. Musical friends and foes: the social cognition of affiliation and control in improvised interactions. *Cognition.* **161**, 94–108. https://doi.org/10.1016/j.cognition.2017.01.019 (2017).

39. Schimmack, U. & Grob, A. Dimensional models of core affect: a quantitative comparison by means of structural equation modeling. *Eur. J. Pers.* **14** (4), 325–345. https://doi.org/10.1002/1099-0984(200007/08)14:4<325::AID-PER380-3.0.CO;2-I (2000).

40. Tan, K. R., Villarino, M. L. & Maderazo, C. Automatic music mood recognition using Russell's two-dimensional valence-arousal space from audio and lyrical data as classified using SVM and Naïve Bayes. *IOP Conf. Series: Mater. Sci. Eng.* **482** (1), 1–6. https://doi.org/10.1088/1757-899X/482/1/012019 (2019).

41. Yang, X., Dong, Y. & Li, J. Review of data features-based music emotion recognition methods. *Multimedia Syst.* **24** (4), 365–389. https://doi.org/10.1007/s00530-017-0559-4 (2018).

42. Juslin, P. N. & Västfjäll, D. Emotional responses to music: the need to consider underlying mechanisms. *Behav. Brain Sci.* **31** (5), 559–575. https://doi.org/10.1017/S0140525X08005293 (2008).

43. Husain, G., Thompson, W. F. & Schellenberg, E. G. Effects of musical tempo and mode on arousal, mood, and spatial abilities. *Music Percept.* **20** (2), 151–171. https://doi.org/10.1525/mp.2002.20.2.151 (2002).

44. Lartillot, O., Eerola, T., Toiviainen, P. & Fornari, J. Multi-feature modeling of pulse clarity: Design, validation and optimization. In J. P. Bello, E. Chew & D. Turnbull (Eds.), ISMIR 2008 international conference on music information retrieval (pp. 521–526). Philadelphia, PA: International Society for Music Information Retrieval. (2008).

45. Gabrielsson, A. Emotion perceived and emotion felt: same and different. *Musicae Sci.* **5**, 123–147. https://doi.org/10.1177/10298649020050S105 (2002).

46. Vieillard, S. et al. Happy, sad, scary and peaceful musical excerpts for research on emotions. *Cogn. Emot.* **22** (4), 720–752. https://doi.org/10.1080/02699930701503567 (2008).

47. Lartillot, O., Toiviainen, P. & Eerola, T. A matlab toolbox for music information retrieval. In (eds Preisach, C., Burkhardt, H., Schmidt-Thieme, L. & Decker, R.) Studies in Classification, data Analysis, and Knowledge Organization (261–268). Springer-Berlin Heidelberg. https://doi.org/10.1007/978-3-540-78246-9_31 (2008).

48. Juslin, P. N. Cue utilization in communication of emotion in music performance. *J. Exp. Psychol. Hum. Percept. Perform.* **26** (6), 1797–1813. https://doi.org/10.1037/0096-1523.26.6.1797 (2000).

49. Tzanetakis, G. & Cook, P. Musical genre classification of audio signals using geometric methods. *IEEE Trans. Speech Audio Process.* **10** (5), 293–302. https://doi.org/10.1109/TSA.2002.800560 (2002).

50. Banchhor, S. K. & Khan, A. Musical instrument recognition using zero crossing rate and short-time energy. *Int. J. Appl. Inform. Syst.* **1** (3), 16–19. https://doi.org/10.5120/ijais12-450131 (2012).

51. McAdams, S. & Giordano, B. L. The perception of musical timbre. In (eds Hallam, S., Cross, I. & Thaut, M. H.) The Oxford Handbook of Music Psychology, (2nd ed., 113–124). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198722946.013.12 (2015).

52. Eerola, T., Lartillot, O. & Toiviainen, P. Prediction of multidimensional emotional ratings in music from audio using multivariate regression models. In K. Hirata, G. Tzanetakis & K. Yoshii (Eds.), Proceedings of the 10th international society for music information retrieval conference (ISMIR 2009) (pp. 621–626). Kobe, Japan: International Society for Music Information Retrieval (2009).

53. Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423. https://doi.org/10.1007/s00530-015-0489-y (1948).

54. Schimmack, U. & Reisenzein, R. Experiencing activation: energetic arousal and tense arousal are not mixtures of valence and activation. *Emotion.* **2** (4), 412–417. https://doi.org/10.1037/1528-3542.2.4.412 (2002).

55. Bradley, M. M. & Lang, P. J. Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry.* **25** (1), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9 (1994).

56. Schäfer, K. & Eerola, T. How listening to music and engagement with other media provide a sense of belonging: an exploratory study of social surrogacy. *Psychol. Music.* **48** (2), 232–251. https://doi.org/10.1177/0305735618795036 (2020).

57. Paquette, S., Peretz, I. & Belin, P. The musical emotional bursts: a validated set of musical affect bursts to investigate auditory affective processing. *Front. Psychol.* **4**, 509. https://doi.org/10.3389/fpsyg.2013.00509 (2013).

58. Savage, P. E. et al. Music as a coevolved system for social bonding. *Behav. Brain Sci.* **44** (e59), 1–22. https://doi.org/10.1017/S0140525X20000333 (2021).

59. Treider, J. M., Kunst, J. R. & Vuoskoski, J. K. The influence of musical parameters and subjective musical ratings on perceptions of culture. *Sci. Rep.* **13** (1), 20682. https://doi.org/10.1038/s41598-023-45805-w (2023).

60. Quinto, L., Thompson, W. F. & Taylor, A. The contributions of compositional structure and performance expression to the communication of emotion in music. *Psychol. Music.* **42** (4), 492–512. https://doi.org/10.1177/0305735613482023 (2013).

Author contributions

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-78156-1.

Correspondence and requests for materials should be addressed to W.F.T.

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

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