International Journal of Psychology

International Journal of Psychology, 2023
Vol. 58, No. 5, 465–475, DOI: 10.1002/ijop.12922

Emotional valence perception in music and subjective arousal: Experimental validation of stimuli

Lena M. Hofbauer and Francisca S. Rodriguez

Research Group Psychosocial Epidemiology and Public Health, German Center for Neurodegenerative Diseases (DZNE), Greifswald, Germany

Musical stimuli are widely used in emotion research and intervention studies. However, reviews have repeatedly noted that a lack of pre-evaluated musical stimuli is stalling progress in our understanding of specific effects of varying music. Musical stimuli vary along a plethora of dimensions. Of particular interest are emotional valence and tempo. Thus, we aimed to evaluate the emotional valence of a set of slow and fast musical stimuli. $N = 102$ (mean age: 39.95, *SD*: 13.60, 61% female) participants rated the perceived emotional valence in 20 fast (>110 beats per minute [bmp]) and 20 slow (<90 bpm) stimuli. Moreover, we collected reports on subjective arousal for each stimulus to explore arousal's association with tempo and valence. Finally, participants completed questionnaires on demographics, mood (profile of mood states), personality (10-item personality index), musical sophistication (Gold-music sophistication index), and sound preferences and hearing habits (sound preference and hearing habits questionnaire). Using mixed-effect model estimates, we identified 19 stimuli that participants rated to have positive valence and 16 stimuli that they rated to have negative valence. Higher age predicted more positive valence ratings across stimuli. Higher tempo and more extreme valence ratings were each associated with higher arousal. Higher educational attainment was also associated with higher arousal reports. Pre-evaluated stimuli can be used in future musical research.

Keywords: Emotional perception; Music perception; Arousal; Individual differences.

Musical stimuli are valuable to researchers interested in emotional processing (Eerola & Vuoskoski, 2012). In addition, the powerful emotional nature of music has led to its therapeutic use, with therapeutic musical interventions being studied with various patient groups, such as people affected by coronary heart disease or people with dementia (Mofredj et al., 2016). However, findings related to music's effects are often conflicting. For instance, as a recent review summarises, it remains unclear how music elicits reward, as biochemical, physical and psychological findings diverge (Reybrouck & Eerola, 2022). In the area of music therapies, meta-analysis shows that observed effects on, for example, cognition, quality of life and anxiety in dementia are too mixed to draw firm conclusions (van der Steen et al., 2018). Reviews converge on the notion that to make sense of such mixed findings, future research should rely on more clearly defined stimuli which should be evaluated in pre-testing (e.g. Eerola & Vuoskoski, 2012; Warrenburg, 2020).

Indeed, a recent review of stimuli used in research focused on music and emotion found that merely 3% of the stimuli reviewed had been pre-evaluated before study use (Warrenburg, 2020). Undeniably, the fact that relevant dimensions along which musical stimuli can be characterised are manifold (including musical, acoustical and emotional features), complicates stimuli pre-evaluation. In general, emotional valence, as the central emotional feature, is considered to be of particular relevance. Yet, describing emotional valence is not trivial. For one thing, one can record the *perceived* emotional valence expressed in the stimulus or the emotion *experienced* while listening to the stimulus. While these correlate, they are not identical, with experienced emotion showing greater variability. For instance, listeners can perceive a musical stimulus to express "sadness" while experiencing a range of mixed

Correspondence should be addressed to Lena M. Hofbauer, Research Group Psychosocial Epidemiology and Public Health, German Center for Neurodegenerative Diseases (DZNE), Ellernholzstr 1-2, 17489 Greifswald, Germany. (E-mail: lena.hofbauer@dzne.de).

We want to thank our colleagues, Sabrina D. Ross and Nadja Ziegert, for their involvement in the stimulus selection process. Further, we want to thank all participants for generously giving their time to this research.

LMH and FSR conceived and designed this study. LMH acquired data and carried out the statistical analysis, which was checked by FSR. Both authors interpreted the data. LMH produced the initial draft of the manuscript, which was further revised by both FSR and LMH. Both authors approved of final submission. The corresponding author attests all authors meet authorship criteria and that no others meeting the criteria have been omitted.

Open Access funding enabled and organized by Projekt DEAL.

emotions, including nostalgia and peacefulness (Taruffi & Koelsch, 2014).

Furthermore, emotional valence can be described using discrete emotion categories or along valence continua (Eerola & Vuoskoski, 2012). In her review, Warrenburg (2020) found that the majority of studies used categorical description with more than half of all stimuli used being described in terms of "sadness," "happiness" and "anger." Yet, valence ratings on continuous scales give more detailed insights, especially when the emotion expressed is ambiguous (Eerola & Vuoskoski, 2011). This is also reflected in greater consistency of continuous as compared to discrete ratings across participants (Song et al., 2016).

Emotional valence perception is directly dependent on acoustical features of music. For instance, dissonance is generally perceived as expressing negative valence, as are pitch numerosity and acoustic roughness (Lahdelma et al., 2022). The most prominently investigated features in this regard, however, are mode and tempo (Parncutt, 2014). Music in major mode is typically perceived as positively valenced, while minor mode is associated with a perceived negative valence (e.g. Morreale et al., 2013). For tempo, high tempo is associated with positive valence while slow tempo is generally perceived as expressing negative valence (e.g. Vieillard et al., 2012). In fact, tempo cues appear powerful enough to outweigh mode when participants judge emotional valence in stimuli with incongruent tempo and mode, at least in non-musicians (Morreale et al., 2013).

Tempo also is of particular interest when investigating therapeutic uses of music, as tempo's effect on nervous system regulation is considered central to therapeutic effects (e.g. Mofredj et al., 2016). However, given the association between tempo and emotional valence, it is difficult to discern whether positive effects can truly be attributed to tempo or whether they relate to the emotional quality of stimuli. Previous studies have typically focused on either tempo (e.g. Liu et al., 2018) or the emotional valence of stimuli (e.g. Proverbio et al., 2015; Song et al., 2016). To identify whether these dimension have independent effects, one would need to assess both *congruent* (fast-positive valence and slow-negative valence) and *incongruent* (slow-positive valence and fast-negative valence) stimuli.

Interpretation of findings is further complicated by the potential role of subjective arousal, that is, the degree of activation felt by a participant in response to a musical stimulus. Higher tempo has been shown to correlate with greater subjective arousal (e.g. Morreale et al., 2013). One study found that when stimuli were matched for subjectively experienced arousal, tempo had no separate effect on physiological parameters (Krabs et al., 2015), suggesting that therapeutic effects attributed to tempo may actually be rooted in varying arousal experiences. Given that high tempo has been associated with both more positively perceived valence and greater arousal experience, one may expect a strong positive correlation between valence and arousal. Correlations have, in fact, been reported (e.g. Bradley & Lang, 2000). However, the relationship remains a subject of debate. While some theoretical models of affect describe arousal and valence as orthogonal (e.g. Posner et al., 2005), converging evidence points towards a V-shaped relationship (Yik et al., 2022). This matches the existence of both high arousal positively valenced music ("happy" or "joyful") music and high arousal, negatively valenced ("angry" or "agitating") music.

Individual differences influencing perception and experience should be considered. For instance, recognition accuracy for negatively valenced emotions has been found to decline with age (Castro & Lima, 2014; Vieillard et al., 2012). Others have not reported an age effect (Song et al., 2016), yet, this may have been due to a more restricted age range. Gender may also have an influence, because women may find it easier to identify emotions expressed (Gabrielsson & Juslin, 1996), although gender effects have not always been found (Castro & Lima, 2014; Song et al., 2016). Moreover, there appears to be a mood-congruence effect in emotional valence perception: participants in a more depressed mood perceived more negative emotion in music, whereas participants with high vigour generally perceived more positive emotion (Vuoskoski & Eerola, 2011). This effect was partially mediated by extraversion (Vuoskoski & Eerola, 2011), indicating an additional role of personality traits.

Moreover, participants' skills and habits may influence the emotional evaluation of music. For example, compared to non-musicians, musicians exhibit more accurate emotion recognition (Castro & Lima, 2014) and perceive fast music as more positively valenced (Liu et al., 2018). These findings may be the result of musicians being more attuned to musical cues. This may well extend to an effect of musical expertise on emotion recognition among non-musicians, however, a previous investigation did not find this (Song et al., 2016). Analytical skills may be an additional determinant, as a meta-analysis showed intelligence to be associated with emotion recognition ability (Schlegel et al., 2020). Besides musical expertise and overall cognitive ability, people differ more specifically in sound preferences and hearing habits. For instance, while some are greatly sensitive to sharp sounds or place high importance on perceiving details in music, others are unaffected by sharp sounds or do not value musical details as highly (Meis et al., 2018). Such differences may modulate attention to valence cues thereby influencing perception.

AIM

The primary aim of this work was to establish a clearly characterised musical stimulus set. To enable future research to differentiate between the effects of emotional valence and tempo, we describe two *congruent* classes of stimuli (i.e. stimuli for which valence and tempo cues match): (a) fast tempo, positive valence, and (b) slow tempo, negative valence, as well as two *incongruent* classes of stimuli: (a) fast tempo, negative valence, and (b) slow tempo, positive valence. We further sought to describe stimuli in terms of subjective arousal felt in response to them, to be able to describe the interrelation between perceived valence, tempo and subjective arousal more clearly. Our secondary aim was to identify individual differences that may be relevant to emotional valence perception and arousal experience.

METHOD

Participants

Participants were recruited via email and advertisement on social media. Non-musicians aged 18 or older were invited to take part in the study. Professional musicians were excluded as their musical perception diverges from that of non-musicians (Castro & Lima, 2014; Liu et al., 2018). Additional exclusion criteria were neuro-cognitive disorders (e.g. dementia and stroke), neurodivergence (e.g., dyslexia or autism), and severe visual/hearing impairments (i.e. impariments severe enough that they might interfere with reading instructions and listening to the music, according to self-report). The final sample consisted of 102 participants (mean age: 39.95, *SD*: 13.6), the majority (60.78%) was female. A full description can be found in Table 1.

TABLE 1
Sample demographics

N	102
Gender, *n* (%)	
Male	38 (37.25)
Female	62 (60.78)
Non-Binary	1 (0.98)
Not reported	1 (0.98)
Age	
Mean (SD)	39.95 (13.6)
Range	19–82
Education, *n* (%)	
Secondary school (Hauptschule)	1 (0.98)
Secondary school (Realschule)	6 (5.88)
Secondary school (Gymnasium)	8 (7.84)
Vocational school	10 (9.8)
University of applied science	15 (14.71)
University	62 (60.78)
POMS, mean (*SD*)	
Dejection/anxiety (POMS-D)	9.96 (11.8)
Fatigue (POMS-F)	20.24 (7.63)
Vigour (POMS-V)	15.87 (8.46)
Anger (POMS-A)	5.67 (6.58)
SP-HHQ, mean (*SD*)	
Annoyance/distraction by background noise (SPHHQ-F1)	21.19 (6.21)
Importance of sound quality (SPHHQ-F2)	6.56 (2.03)
Noise sensitivity (SPHHQ-F3)	9 (3.76)
Avoidance of unpredictable sounds (SPHHQ-F4)	7.19 (2.81)
Openness towards loud/new sounds (SPHHQ-F5)	10.61 (2.51)
Preferences for warm sounds (SPHHQ-F6)	7.36 (1.78)
Details of environmental sounds/music (SPHHQ-F7)	7.63 (1.58)
Gold-MSI, mean (*SD*)	
General factor of musical sophistication (GMSI-General)	68.13 (13.65)
TIPI, mean (*SD*)	
Openness to experience (TIPI-O)	10.85 (2.01)
Conscientiousness (TIPI-C)	10.9 (2.21)
Extraversion (TIPI-E)	8.82 (2.49)
Agreeableness (TIPI-A)	10.25 (1.88)
Emotional stability (TIPI-STAB)	9.91 (2.21)

Note: Descriptive statistics of demographic information and individual difference scores.

Stimuli

Stimuli were 30 seconds-long excerpts from a total of 40 music recordings. The excerpts were created in Audacity by selecting the 30 seconds around the midpoint of each original audio file. We chose the exact mid-section rather than having researchers choose, e.g. according to perceived representativeness, to avoid the introduction of bias. To ensure that pieces were comparable, criteria for selection were that they (a) be entirely instrumental (i.e. no human voices or environmental sounds) and (b) be performed by more than two instruments. Half of excerpts were taken from classical musical pieces (i.e. broadly Western art music), half from movie soundtracks. We chose classical music, on the one hand, as this is the genre used by the greatest proportion of music emotion publications, and has the additional advantage of pieces often being strictly instrumental. Movie soundtracks, on the other hand, have more recently gained traction in music emotion research as they are specifically composed with the expression of emotion in mind (Warrenburg, 2020). To make the stimuli selection reproducible, we have added a detailed description of the excerpts in Table S1.

Low tempo was defined as a bpm $<$ 90 (range: 59–87 bpm) and high tempo as a bpm $>$ 110 (range: 116–184 bpm). Three blinded researchers selected stimuli they expected to be perceived as either having positive or negative emotional valence. The goal was to assign 10 stimuli to each of four target stimuli classes: (a) fast tempo, expected positive valence, (b) slow tempo, expected positive valence, (c) fast tempo, expected negative valence, (d) slow tempo, expected negative valence. This was an attempt to ensure that stimuli presented were sufficiently varied; we do not take into account which valence was expected in analyses. The tempo of the pieces was identified using “Tempo-CNN” in Python.

Questionnaires on individual differences

Demographic information questionnaire

Participants were asked to report their age. They also reported their gender (male, female, non-binary), except one participant, who did not report their gender. Further, participants reported their highest academic degree according to the German school system (secondary school [Hauptschule], secondary school [Realschule], secondary school [Gymnasium], vocational school, University of Applied Science, University); this was used as a proxy for intelligence (Ritchie & Tucker-Drob, 2018).

Profile of mood states (POMS)

We used the German version of the POMS (see Albani et al., 2005), assessing the current mood of participants using 35 items. Participants rate each item in terms of how strongly it applied to their current mood on a Likert Scale ranging from 0 (“Not at all”) to 6 (“Extremely”). The items form four subscales, subscale scores were calculated by summing the score of the relevant items: dejection/anxiety (*POMS-D*, 14 items, range: 0–84), vigour (*POMS-V*, 7 items, range: 0–42), anger (*POMS-A*, 7 items, range: 0–42) and fatigue (POMS-F, 7 items, range: 0–42). According to the manual, one missing item per subscale was tolerated and replaced by the mean of the remaining items of this scale. This rule had to be applied in only one case.

Sound preference and hearing habits questionnaire (SP-HHQ)

The German SP-HHQ (Meis et al., 2018) uses 23 items to assess sound preferences and hearing habits. Items are rated on a scale from 1 (“Does not apply at all”) to 5 (“Applies fully”) and summed to give seven factor scores: Annoyance/distraction by background noise (*SPHHQ-F1*, 8 items, range: 8–40), Importance of sound quality (*SPHHQ-F2*, 4 items, range: 4–20), Noise sensitivity (*SPHHQ-F3*, 3 items, range: 3–15), Avoidance of unpredictable sounds (*SPHHQ-F4*, 3 items, range: 3–15), Openness towards loud/new sounds (*SPHHQ-F5*, 3 items, range: 3–15), Preferences for warm sounds (*SPHHQ-F6*, 2 items, range 2–10), Details of environmental sounds/music (*SPHHQ-F7*, 2 items, range 2–10),

Gold music sophistication index (Gold-MSI)

The Gold-MSI assesses musical attitudes, behaviours, and skills. It has been translated into German, and its subscales can be used independently (Schaal et al., 2014). The 18-item General Factor of Musical Sophistication Subscale (*GMSI-General*) for a general overview over participants’ musicality. The majority of items are rated on a scale from 1 (“Completely Disagree”) to 7 (“Completely Agree”). Three items use answer options corresponding to scores of 1–7 instead of a Likert scale. The sum of all item scores gives the total score (range: 18–126).

Ten-item personality index (TIPI)

We used the German Version of the TIPI (Muck et al., 2007). It assesses the “Big 5” personality domains: Openness to Experience (*TIPI-O*), Conscientiousness (*TIPI-C*), Extraversion (*TIPI-E*), Agreeableness (*TIPI-A*) and Emotional Stability (*TIPI-STAB*). Two items describe each of the domains, one of these always being reverse coded. The items’ scale ranges from 1 (“Disagree strongly”) to 7 (“Agree strongly”), and items of a subscale are summed (range of each subscale: 2–14).

Procedure

The study was realised in form of an online survey generated using SoSci Survey and hosted on their server s2survey.net. Participants were first presented with study information and an online consent screen and could provide consent by double opt-in. That is, participants entered their email address and actively confirmed that they consented to participation by clicking a confirmation link contained in an email. The link directed them to the survey. In part one of the survey, participants completed a sequence of questionnaires on individual differences. In the second part, participants were asked to rate the musical excerpts on two continuous sliding scales. First, in terms of perceived emotional valence, ranging from “very negative” to “very positive” (Screen display: “*Please indicate whether you perceive the emotion expressed in the music to be negative or positive by moving the slider accordingly*.”). Second, in terms of subjective arousal, ranging from “not at all arousing” to “very arousing” (Screen display: “*Please also indicate how emotionally arousing you find the music by moving the slider accordingly*.”). Scale positions were translated into values between −50 and 50 for analysis. A score of 0 represents the neutral scale position. The music stimuli were presented in groups of four per survey page. Groups were generated by a random generator and kept constant. The order of in which pages were presented was re-shuffled by a random generator for each participant.

Analysis

Data analysis was conducted using R (Version 4.0.2) in RStudio (Version 1.2.5042). We used a significance level of $p < .05$.

Emotional valence ratings

To determine the perceived emotional valence of stimuli we estimated the intercept for each stimulus by fitting mixed-effect models with a zero intercept, stimulus as fixed effect, and a random by-participant intercept (R package "lme4"). We considered stimuli with estimates and confidence intervals (CIs) that were exclusively positive as having positive valence. Stimuli with negative estimates and CIs that were exclusively negative were considered as having negative valence. If the CI included zero, valence could not be determined. By using these estimates we avoided aggregating rating data and thus losing substantial information. We also calculated the skewness of perceived valence ratings for each stimulus (R package "e1071"). Skewness expresses to which degree the distribution of ratings diverges from a symmetrical distribution. A symmetrical distribution would suggest no tendency, that is, neither predominantly positive nor predominantly negative valence ratings. Positive skew would indicate a right-skew, while negative values would indicate a left-skew. Absolute skews of $>.5$ would indicate a moderate tendency and absolute skews of >1 would indicate a strong tendency. Therefore, higher absolute skews would indicate less ambiguity in the perceived emotional valence of a stimulus.

Arousal ratings

To determine how arousing participants found stimuli to be, we calculated arousal estimates by using a mixed-effect model with a zero intercept, stimulus as fixed effect, and a random by-participant intercept. If estimate and CI were positive, arousal was labelled as high; if estimate and CI were negative, arousal was considered low. Additionally, we considered skewness of arousal ratings to get an indication of ambiguity.

Further, we separately investigated the relationship of arousal ratings with two predictors of interest: (a) emotional valence rating, and (b) tempo. More specifically, we investigated whether these predictors had linear or quadratic relationships with arousal ratings by fitting both linear and quadratic models (R package "stats") and comparing model fit. To this end, we used the Aikake Information Criterion (AIC) and the Bayesian Information Criterion (BIC).

Individual differences

In the absence of specific predictions regarding the effects of individual difference dimensions, we calculated exploratory mixed-effect models of perceived valence and subjective arousal including the individual difference dimension collected via questionnaires and by-participant intercepts. Inclusion of interaction terms was not feasible, as this induced problematic levels of multicollinearity, as indicated by variance inflation factors >5 (R package "performance"). To prevent undue influence of extreme scorers on the results of this analysis, we further identified *SD*-based outliers. We identified one participant who was more than three *SD*s older than the average age, three whose POMS-D scores were more than three *SD*s above the mean, and three whose scores on the SPHHQ-F6 were more than 3 *SD*s below the mean. Removing outliers one-by-one and comparing results to the full sample, none were identified as influential. Thus, we did not exclude them from analyses. To avoid group sizes of $n=1$, we removed the non-binary participant, the participant who did not report their gender, and the only one participant who had the education status secondary school (Hauptschule) (see Table 1). As a result, data of $n=99$ participants were included in the analyses of individual differences.

RESULTS

Emotional valence ratings

Based on mixed-model estimates and CIs, we labelled 19 stimuli as having positive valence. Of these, 11 were fast tempo stimuli (IDs: 05, 06, 07, 09, 10, 25, 26, 27, 28, 29, 30; see Table 2), with skews varying between −0.05 and −1.89. Thus, skews indicated a weak to strong tendency towards positive ratings. Eight of the positively valenced stimuli had a slow tempo (IDs: 16, 17, 18, 20, 36, 38, 39, 40), with skews varying between 0.21 and −0.47 (see Table 2 and Table S2). That is, skews indicated weak tendencies in both positive and negative directions. In addition, 16 stimuli were labelled as having negative valence, half with slow tempo (IDs: 11, 12, 14, 19, 32, 33, 34, 35; skew range: 0.34–1.13) and half with fast tempo (IDs: 01, 02, 03, 04, 21, 22, 23, 24; skew range: 0.36–1.42). In other words, ratings were slightly-to-strongly biased towards negative valence. Five stimuli could not be classified as having positive or negative valence.

Arousal ratings

Twenty-one stimuli were reported to induce high subjective arousal. Of these, 10 were fast with positive valence (IDs: 05, 06, 07, 08, 09, 10, 25, 26, 27, 28, 29; skew range: −1.25 to −0.30; Table 2), 7 were fast with negative valence (IDs: 01, 02, 03, 21, 22, 23, 24; skew range: −0.95 to −0.35) and 3 were slow with negative valence (IDs: 14, 19, 34, skew range: −0.43 to −0.04). That is, arousal ratings were overall slightly-to-strongly biased towards higher arousal for fast stimuli and only slightly biased towards high arousal for slow negative stimuli. One stimulus with slow tempo and positive valence induced low arousal (ID: 37). Skew for the latter was zero, indicating

TABLE 2
Description of stimuli

Musical piece (ID number, name, artist)	*Tempo (bpm, Label[a])*	*Perceived emotional valence, estimate (95% CI[b]), Label[c]*	*Per. emo. valence skew*	*Subjective arousal, estimate (95% CI[b]), Label[d]*	*Subjective arousal skew*
01, Eine Nacht Auf Dem Kahlen Berge *by Modest Mussorgsky*	116, *Fast*	−25.35 (−29.31, −21.35), *Negative Valence*	1.42	7.87 (2.82, 13.03), *High Arousal*	−0.46
02, Le Sacre du Printemps, Part 2: The Sacrifice: Sacrificial Dance *by Igor Stravinsky*	128, *Fast*	−18.14 (−22.06, −14.06), *Negative Valence*	1.41	9.62 (4.39, 14.75), *High Arousal*	−0.60
03, Scythian Suite, Op.20—"Ala And Lolly": 2. The Enemy God And The Dance Of The Spirits Of Darkness *by Sergei Prokofiev*	118, *Fast*	−11.2 (−15.36, −7.24), *Negative Valence*	0.91	6.95 (1.07, 11.75), *High Arousal*	−0.43
04, The Planets: I. Mars, The Bringer of War *by Gustav Holst*	135, *Fast*	−15.31 (−19.59, −11.49), *Negative Valence*	1.14	1.89 (−3.55, 6.69)	−0.16
05, Romeo and Juliet, Op.64—Act 1: The Duke's Command—Interlude, *by Sergei Prokofiev*	125, *Fast*	8.21 (4.03, 12.47), *Positive Valence*	−0.12	15.84 (10.77, 21.09), *High Arousal*	−0.77
06, Der Zigeunerbaron, Act III: "March. Hurrah, die Schlacht" *by Johann Strauss II*	118, *Fast*	33.23 (29.10, 37.33), *Positive Valence*	−1.89	19.69 (1.67, 12.10), *High Arousal*	−0.42
07, Eine kleine Nachtmusik K. 525: Allegro *by Wolfgang Amadeus Mozart*	130, *Fast*	35.44 (31.71, 39.57), *Positive Valence*	−1.88	19.86 (14.35, 24.80), *High Arousal*	−0.84
08, Kammermusik No. 1, Op. 24 No. 1: I. Sehr schnell und wild, *by Paul Hindemith*	139, *Fast*	0.52 (−3.07, 4.87)	0.14	14.55 (9.57, 20.43), *High Arousal*	−0.79
09, Keyboard Concerto No. 3 in D Major, BWV 1054: III. Allegro *by Johann Sebastian Bach*	139, *Fast*	29.87 (25.86, 34.14), *Positive Valence*	−0.93	10.82 (5.97, 15.51), *High Arousal*	−0.45
10, William Tell Overture *by Gioachino Rossini*	167, *Fast*	29.58 (25.51, 33.64), *Positive Valence*	−1.22	25.60 (20.13, 30.37), *High Arousal*	−1.25
11, Cantus in memoriam Benjamin Britten, *by Avo Pärt*	87, *Slow*	−4.70 (−8.91, −0.35), *Negative Valence*	0.34	−0.52 (−5.56, 4.50)	−0.33
12, Dido and Aeneas, Z. 626 / Act 3: "When I Am Laid In Earth" Dido's Lamento *by Henry Purcell*	87, *Slow*	−12.57 (−16.93, −8.62), *Negative Valence*	1.10	3.96 (−0.86, 9.07)	−0.06
13, Fratres For String Orchestra And Percussion, *by Avo Pärt*	87, *Slow*	−2.80 (−6.56, 1.29)	0.82	−1.75 (−6.46, 3.26)	−0.17
14, (La) Mélancolie *by Ole Bornemann Bull*	87, *Slow*	−10.64 (−14.76, −7.04), *Negative Valence*	0.94	8.18 (2.89, 13.71), *High Arousal*	−0.43
15, Peer Gynt Suite No.1, Op.46–2. The Death Of Aase, *by Edvard Grieg*	87, *Slow*	−0.64 (−4.63, 3.18)	0.52	−2.97 (−8.36, 2.29)	−0.05
16, Carmen Suite No. 1: III. Intermezzo, *by Georges Bizet*	71, *Slow*	6.48 (2.32, 10.21), *Positive Valence*	0.21	−3.31 (−8.91, 1.22)	0.12
17, Piano Concerto No. 5 in E-Flat Major, Op. 73: II. Adagio un poco mosso *by Ludwig van Beethoven*	87, *Slow*	15.44 (11.83, 19.05), *Positive Valence*	−0.38	−2.45 (−7.75, 2.59)	0.11
18, Viola Concerto in G Major: Largo, *by Georg Philipp Telemann*	87, *Slow*	6.05 (2.31, 10.21), *Positive Valence*	0.06	0.80 (−4.09, 6.12)	−0.21
19, Concerto Grosso in C Major, HWV 318, Alexander's Feast: Largo, *by George Frideric Handel*	87, *Slow*	−5.50 (−9.81, −1.52), *Negative Valence*	0.42	5.36 (0.69, 10.39), *High Arousal*	−0.04
20, Oboe Concerto in B-Flat Major: Largo e mesto, *by Carl Philipp Emanuel Bach*	87, *Slow*	5.20 (1.18, 9.42), *Positive Valence*	0.08	−4.23 (−9.41, 1.15)	−0.06
21, The Droid *by Jerry Goldsmith*	130, *Fast*	24.58 (−28.28, −21.01), *Negative Valence*	1.21	6.92 (1.59, 11.79), *High Arousal*	−0.35

TABLE 2
Continued

Musical piece (ID number, name, artist)	*Tempo (bpm, Label[a])*	*Perceived emotional valence, estimate (95% CI[b]), Label[c]*	*Per. emo. valence skew*	*Subjective arousal, estimate (95% CI[b]), Label[d]*	*Subjective arousal skew*
22, Assets And Targets *by John Powell*	116, *Fast*	−7.75 (−11.72, −4.02), *Negative Valence*	0.67	13.15 (7.80, 18.44), *High Arousal*	−0.47
23, Chasing Luisa Rey, *by Tom Tykwer, Johnny Klimek, Reinhold Heil*	130, *Fast*	−4.02 (−7.63, −0.15), *Negative Valence*	0.36	19.10 (14.23, 24.24), *High Arousal*	−0.95
24, Impulse, *by Hans Zimmer*	120, *Fast*	−18.98 (−22.92, −15.44), *Negative Valence*	0.68	11.86 (6.66, 17.84), *High Arousal*	−0.61
25, Mombasa, *by Hans Zimmer*	145, *Fast*	5.27 (1.65, 9.47), *Positive Valence*	−0.05	13.99 (9.17, 19.56), *High Arousal*	−0.77
26, Little Women, *by Alexandre Desplat*	156, *Fast*	6.23 (2.07, 9.98), *Positive Valence*	0.09	7.71 (2.07, 12.56), *High Arousal*	−0.30
27, Making Water *by Harry Gregson-Williams*	116, *Fast*	19.07 (15.00, 23.22), *Positive Valence*	−0.18	8.33 (3.34, 13.32), *High Arousal*	−0.50
28, Superman Returns: Main Theme *by John Ottman*	120, *Fast*	27.27 (23.25, 30.88), *Positive Valence*	−1.07	14.22 (9.36, 19.06), *High Arousal*	−0.63
29, The Return of the Eagle *by Atli Örvarsson*	130, *Fast*	23.59 (19.65, 27.35), *Positive Valence*	−1.07	16.80 (11.79, 21.82), *High Arousal*	−0.85
30, Waltz to the Death *by Danny Elfmann*	184, *Fast*	23.87 (19.81, 27.79), *Positive Valence*	−0.83	0.07 (−4.52, 5.20)	−0.19
31, Atonement, *by Dario Marianelli*	87, *Slow*	−3.49 (−7.54, 0.48)	0.38	0.28 (−4.99, 5.63)	0.09
32, Long, Long Time Ago *by Javier Navarrete*	84, *Slow*	−13.38 (−17.30, −9.47), *Negative Valence*	0.95	0.82 (−4.89, 6.45)	−0.07
33, Pan's Labyrinth Lullaby *by Javier Navarette*	73, *Slow*	−9.63 (−13.68, −5.47), *Negative Valence*	0.83	−0.77 (−5.78, 3.87)	−0.28
34, Solomon, *by Hans Zimmer*	87, *Slow*	−7.08 (−11.09, −3.07), *Negative Valence*	0.60	9.50 (4.21, 14.33), *High Arousal*	−0.39
35, The Revenant: Main Theme *by Ryuichi Sakamoto*	87, *Slow*	−18.20 (−21.79, −14.24), *Negative Valence*	1.13	−8.28 (−13.41, −3.05)	0.10
36, Dawn *by Dario Marianelli*	81, *Slow*	19.07 (15.71, 23.02), *Positive Valence*	−0.32	2.19 (−2.39, 7.22)	−0.15
37, Rooftop Kiss, *by James Horner*	73, *Slow*	1.11 (−2.52, 5.17)	0.29	−7.71 (−12.80, −2.58), *Low arousal*	0.00
38, Bogart & Bergman *by Justin Hurwith*	59, *Slow*	23.39 (18.77, 27.16), *Positive Valence*	−0.42	−11.24 (−16.42, −5.81)	0.21
39, Love Theme for Nata *by Ennio Morricone*	87, *Slow*	15.14 (11.19, 19.09), *Positive Valence*	−0.16	−0.10 (−4.75, 4.74)	−0.19
40, First Youth *by Ennio Morricone*	67, *Slow*	17.61 (13.80, 21.28), *Positive Valence*	−0.47	−4.08 (−8.93, 0.88)	0.08

Note: CI = confidence interval; Emo. = emotional; Per. = perceived; Subj. = subjective. [a]Tempo was labelled as "Fast" if >110 bpm, as "Slow" if <90 bpm. [b]CIs were calculated using non-parametric bootstrapping. [c]Perceived valence was labelled as "Positive Valence" if estimate and CI were exclusively positive, as "Negative Valence" if estimate and CI were exclusively negative. [d]Subjective arousal was labelled as "High arousal" if estimate and CI were exclusively positive, as "Low Arousal" if estimate and CI were exclusively negative.

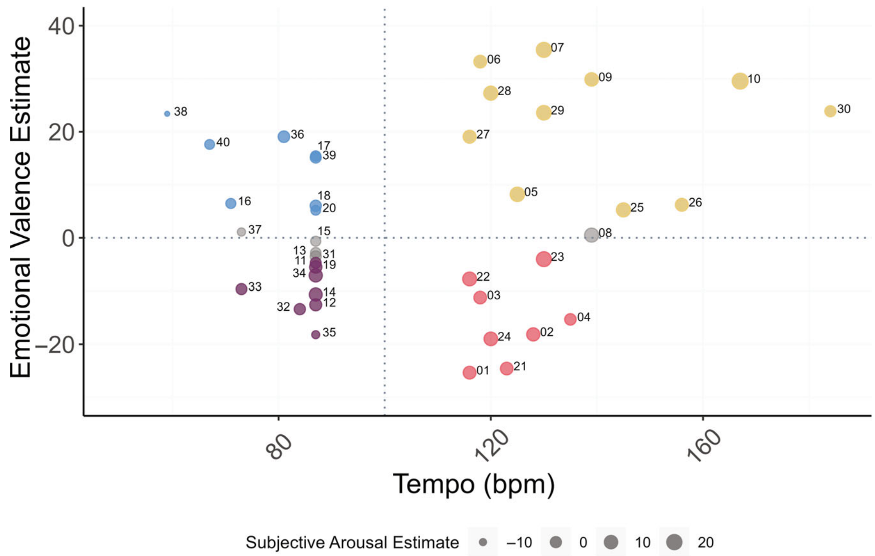

Figure 1. Illustration of stimuli according to tempo in beats per minute (bpm), perceived emotional valence and subjective arousal. The dotted grid signifies the neutral value for the arousal rating (*y*-axis) and the mid-point between fast (>110 bpm) and slow (<90 bpm) tempo (*x*-axis). Points mark the perceived valence estimate and are scaled according to arousal estimates. Points are also labelled with their ID Number and coloured according tempo and valence. Violet: slow, negative valence; blue: fast, negative valence; red: fast, negative valence; yellow: fast, positive valence; grey: no valence could be determined. For more details see Table 2.

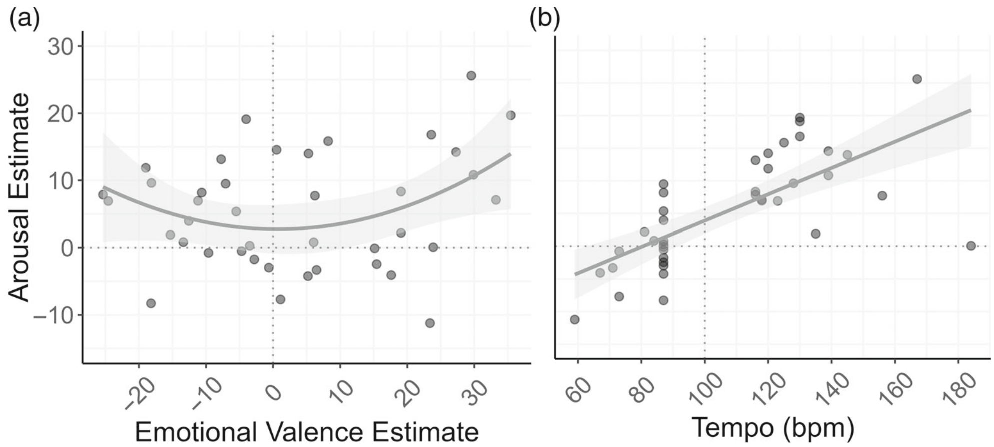

Figure 2. Illustration of arousal ratings in relation to perceived emotional valence and tempo. (a) Shows the relation between arousal ratings and perceived emotional valence. The dotted grid signifies the neutral value for the arousal rating (*y*-axis) and for the perceived emotional valence (*x*-axis). (b) Shows the relation between arousal ratings and tempo. The dotted grid signifies the neutral value for the arousal rating (*y*-axis) and the mid-point between fast (>110 bpm) and slow (<90 bpm) tempo (*x*-axis). The dark grey lines are fitted regression lines; the light grey shading represents the 95% confidence intervals (CIs).

that overall ratings showed no tendency towards higher or lower arousal. Figure 1 illustrates how stimuli vary in terms of tempo, perceived valence and subjective arousal.

Modelling the relationship between emotional valence ratings and arousal ratings, the quadratic model fit outperformed the linear model fit on AIC (38,222 vs. 38,380) and BIC (38,247 vs. 38,399). The significant linear term indicated that arousal slightly increased with valence ($\beta = .06$, $p < .001$), while the quadratic term showed that the slope of this association increased further with increasingly positive valence ratings ($\beta = .01$, $p < .001$; see Figure 2a). In addition, we found a linear association between tempo and arousal. While the quadratic model had a slightly smaller AIC than the linear model (38,230 vs. 38,236), the BIC was equal. The linear term for tempo showed that increasing tempo was associated with higher arousal ratings ($\beta = .20$, $p < .001$; see Figure 2b).

Individual differences

Predicting the perceived emotional valence across stimuli, the only significant predictor was age ($\beta = .20$, $p = .011$), indicating that with increasing age participants generally perceived stimuli's valence as more positive. The sole significant predictor of subjective arousal was education ($\beta = 2.54$, $p = 0.025$), indicating that higher levels of education were associated with reports of greater arousal across stimuli.

DISCUSSION

The primary aim of this study was to evaluate musical stimuli of slow and fast tempo in terms of emotional valence. Using the emotional valence perceptions of 102 participants, we calculated perceived valence estimates for each stimulus and confidence intervals surrounding these estimates. Based on these, we were able to identify 19 congruent (11 fast, positively valenced, 8 slow, negatively valenced) and 16 incongruent (8 slow, positively valenced, 8 fast, negatively valenced) stimuli. Inspection the distribution skews of ratings indicated that slow, positive valenced stimuli were often ambiguous to participants, that is, skews were comparatively small. This is in line with a previously noted high potential for confusion between "tenderness" and "sadness," that is, a positively and negatively valenced emotion, respectively, that are both mostly expressed through slow tempo (Eerola & Vuoskoski, 2011). Researchers using slow, positively valenced stimuli should thus be aware that the stimuli may be perceived as ambiguous, which may result in mixed emotional responses.

Secondly, we estimated subjective arousal, finding 21 stimuli to induce high arousal and one stimulus to induce low arousal. That only 22/40 stimuli (55%) could be clearly described in terms of arousal reflects the overall lower agreement between participants, which can also be observed when inspecting rating distributions and their skew. Overall more mixed subjective arousal experiences matches findings showing greater variability in subjective music experiences than in music perception (Song et al., 2016). We found the most extreme skews in arousal ratings for fast, positive stimuli and observed a linear relationship between tempo and arousal, so that higher tempo related to higher arousal ratings. This supports previous reports of such a positive association between tempo and arousal (Morreale et al., 2013). The relationship between valence and arousal, however, was not linear. Instead, we observe a quadratic relationship: Highly negative and highly positive valence, in particular, were associated with greater subjective arousal. This suggests an asymmetrical V-shaped relationship, which is indeed globally the most commonly observed pattern (Yik et al., 2022). This V-shape may provide a unifying explanation for otherwise conflicting results: both "joyful" (fast, positive) and "touching" (slow, negative) stimuli have been found to stimulate cognition (Proverbio et al., 2015). If both were characterised by strong emotional valence and arousal, this combination of findings could be explained via a shared arousal-cognitive stimulation route.

Investigating the role of individual differences in overall valence perception across stimuli, the only significant predictor was age, with greater age relating to more positive valence perception. This finding echoes previous research on a "positivity bias" with age, that is, a tendency in older adults to perceive stimuli as more positively valenced (Castro & Lima, 2014; Vieillard et al., 2012). Such a tendency appears to be rooted in a bias towards processing positive information. This is believed to be due to older adults prioritising goals relating to emotional meaning and satisfaction, meaning that their focus in processing is actively shifted towards positive information (Reed et al., 2014). It is also possible that hearing difficulties, which are increasingly common with higher age, influenced the perception of musical valence clues. While we did exclude participants with self-reported severe hearing difficulty, low-grade impairments may have been present. However, we do not believe it likely such impairments played a major role in the positivity effect observed: First, it has been shown that those with hearing loss perceive pleasant sounds as less pleasant than those without (Picou et al., 2022). Therefore, if anything, the positivity may be attenuated if individuals with hearing loss participated. Moreover, the difference between those with and without hearing loss was found to be reduced in an online setting versus a laboratory setting (Picou et al., 2022). Finally, the age "positivity bias" is present across modalities (Reed et al., 2014), thus, it appears unlikely that it is attributable to a single sensory impairment.

For subjective arousal rating, the only significant predictor was education, with higher educational attainment being associated with higher arousal ratings. In previous research, those with higher cognitive ability have been shown to have better emotion recognition ability (Schlegel et al., 2020). It is conceivable that greater attunement to valence cues resulted in greater felt arousal. Yet, to our knowledge, there is no direct evidence of such an interaction between valence cues, cognitive ability on arousal. Another potential explanation is that those with higher education may have been more exposed to music similar to that presented in the experiment and that greater familiarity resulted in greater subjective arousal, as has been observed in the literature (Van Den Bosch et al., 2013). Further investigation is warranted. Overall, it should be noted that the results of exploratory models on individual differences are only indicative; future research would need to evaluate more specifically how such individual dimensions influence musical perception and experience and under what circumstances these difference affect outcomes of interest.

An advantage of the research is that we were able to recruit a large sample, which, compared to studies recruiting from a pool of university students (e.g.

Eerola & Vuoskoski, 2011; Song et al., 2016), was more diverse in education and age. Yet, there also were notable limitations. One, to limit the duration of the survey procedure, we used 30-second excerpts, not the entire musical pieces. However, it has been observed that emotion can be reliably perceived in this time (Warrenburg, 2020). Moreover, while we accounted for sound preferences and hearing habits, we did not directly account for music preference or liking, which might have influenced arousal ratings. However, music preferences appear unrelated to emotional judgements (Vieillard et al., 2012) and as such should not have influenced the central results of this work. Further, as this was not a controlled experiment under laboratory conditions, we cannot discount the possibility that there were other sources of variability. For one, while participants were provided with contact details, none contacted us. Thus, it is possible that participants did not get clarification in situations where the wording of questions was ambiguous to them. For instance, while we intended to get ratings on perceived valence and experienced arousal, some may have misinterpreted the wording of items. Two, as we could not control the listening situation, background noises and other distractions may have affected results. However, while valence ratings made by individuals without self-reported hearing loss are less extreme when made online compared to when made offline, the overall pattern should not be affected (Picou et al., 2022). Finally, it should be noted that the analyses of individual differences need to be interpreted with caution as lacking power or range restriction may affect results.

In conclusion, we were able to identify a stimulus set of 35 stimuli of both slow and fast tempo for which participants' evaluations of emotional valence met our criteria. We hope that the availability of such a pre-evaluated stimulus set will aid future investigations in clarifying the effect of tempo and emotional valence expressed in music. In addition, we recorded arousal ratings and showed significant, yet small, positive associations between these ratings and tempo. Further, we found a quadratic association between valence and arousal ratings, so that more extreme emotional valence, be it positive or negative, was associated with greater subjective arousal. Finally, we identified individual difference that influence both perceived emotional valence and arousal ratings. These will be important to consider in future research.

COMPLIANCE WITH ETHICAL STANDARDS

All procedures performed in studies involving human participants were in accordance with the ethical standards of the University Medicine Greifswald's ethics committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all participants included in the study.

Manuscript received April 2022
Revised manuscript accepted April 2023
First published online May 2023

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Table detailing the origin of musical excerpts used.

Table S2. Table showing the distributions of emotional valence ratings and subjective arousal ratings for each of the musical stimuli.

REFERENCES

Albani, C., Blaser, G., Geyer, M., Schmutzer, G., Brähler, E., Bailer, H., & Grulke, N. (2005). The German short version of "Profile of Mood States" (POMS): Psychometric evaluation in a representative sample. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, *55*(7), 324–330.

Bradley, M. M., & Lang, P. J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology*, *37*(2), 204–215.

Castro, S. L., & Lima, C. F. (2014). Age and musical expertise influence emotion recognition in music. *Music Perception: An Interdisciplinary Journal*, *32*(2), 125–142.

Eerola, T., & Vuoskoski, J. K. (2011). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, *39*(1), 18–49.

Eerola, T., & Vuoskoski, J. K. (2012). A review of music and emotion studies: Approaches, emotion models, and stimuli. *Music Perception: An Interdisciplinary Journal*, *30*(3), 307–340.

Gabrielsson, A., & Juslin, P. N. (1996). Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of Music*, *24*(1), 68–91.

Krabs, R. U., Enk, R., Teich, N., & Koelsch, S. (2015). Autonomic effects of music in health and Crohn's disease: The impact of isochronicity, emotional valence, and tempo. *PLoS One*, *10*(5), e0126224.

Lahdelma, I., Armitage, J., & Eerola, T. (2022). Affective priming with musical chords is influenced by pitch numerosity. *Musicae Scientiae*, *26*(1), 208–217. https://doi.org/10.1177/1029864920911127

Liu, Y., Liu, G., Wei, D., Li, Q., Yuan, G., Wu, S., Wang, G., & Zhao, X. (2018). Effects of musical tempo on musicians' and non-musicians' emotional experience when listening to music. *Frontiers in Psychology*, *9*, 2118. https://doi.org/10.3389/fpsyg.2018.02118

Meis, M., Huber, R., Fischer, R.-L., Schulte, M., Spilski, J., & Meister, H. (2018). Development and psychometric properties of the sound preference and hearing habits questionnaire (SP-HHQ). *International Journal of Audiology*, *57*(suppl 3), S118–S129.

Mofredj, A., Alaya, S., Tassaioust, K., Bahloul, H., & Mrabet, A. (2016). Music therapy, a review of the potential therapeutic benefits for the critically ill. *Journal of Critical Care*, *35*, 195–199. https://doi.org/10.1016/j.jcrc.2016.05.021

Morreale, F., Masu, R., De Angeli, A., & Fava, P. (2013). *The effect of expertise in evaluating emotions in music*. The 3rd International Conference on Music & Emotion, Jyväskylä, Finland, June 11–15, 2013.

Muck, P. M., Hell, B., & Gosling, S. D. (2007). Construct validation of a short five-factor model instrument. *European Journal of Psychological Assessment*, *23*(3), 166–175.

Parncutt, R. (2014). The emotional connotations of major versus minor tonality: One or more origins? *Musicae Scientiae*, *18*(3), 324–353.

Picou, E. M., Singh, G., & Russo, F. A. (2022). A comparison between a remote testing and a laboratory test setting for evaluating emotional responses to non-speech sounds. *International Journal of Audiology*, *61*(10), 799–808. https://doi.org/10.1080/14992027.2021.2007422

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, *17*(3), 715–734.

Proverbio, A. M., Nasi, V. L., Arcari, L. A., De Benedetto, F., Guardamagna, M., Gazzola, M., & Zani, A. (2015). The effect of background music on episodic memory and autonomic responses: Listening to emotionally touching music enhances facial memory capacity. *Scientific Reports*, *5*, 15219.

Reed, A. E., Chan, L., & Mikels, J. A. (2014). Meta-analysis of the age-related positivity effect: Age differences in preferences for positive over negative information. *Psychology and Aging*, *29*(1), 1–15.

Reybrouck, M., & Eerola, T. (2022). Musical enjoyment and reward: From hedonic pleasure to eudaimonic listening. *Behavioral Sciences*, *12*(5), 154.

Ritchie, S. J., & Tucker-Drob, E. M. (2018). How much does education improve intelligence? A meta-analysis. *Psychological Science*, *29*(8), 1358–1369. https://doi.org/10.1177/0956797618774253

Schaal, N. K., Bauer, A.-K. R., & Müllensiefen, D. (2014). Der Gold-MSI: replikation und validierung eines fragebogeninstrumentes zur messung musikalischer erfahrenheit anhand einer deutschen stichprobe. *Musicae Scientiae*, *18*(4), 423–447.

Schlegel, K., Palese, T., Mast, M. S., Rammsayer, T. H., Hall, J. A., & Murphy, N. A. (2020). A meta-analysis of the relationship between emotion recognition ability and intelligence. *Cognition and Emotion*, *34*(2), 329–351. https://doi.org/10.1080/02699931.2019.1632801

Song, Y., Dixon, S., Pearce, M. T., & Halpern, A. R. (2016). Perceived and induced emotion responses to popular music: Categorical and dimensional models. *Music Perception: An Interdisciplinary Journal*, *33*(4), 472–492.

Taruffi, L., & Koelsch, S. (2014). The paradox of music-evoked sadness: An online survey. *PLoS One*, *9*(10), e110490.

Van Den Bosch, I., Salimpoor, V., & Zatorre, R. (2013). Familiarity mediates the relationship between emotional arousal and pleasure during music listening. *Frontiers in Human Neuroscience*, *7*, 534. https://doi.org/10.3389/fnhum.2013.00534

van der Steen, J. T., Smaling, H. J., van der Wouden, J. C., Bruinsma, M. S., Scholten, R. J., & Vink, A. C. (2018). Music-based therapeutic interventions for people with dementia. *Cochrane Database of Systematic Reviews*, *2018*(7), CD003477.

Vieillard, S., Didierjean, A., & Maquestiaux, F. (2012). Changes in the perception and the psychological structure of musical emotions with advancing age. *Experimental Aging Research*, *38*(4), 422–441.

Vuoskoski, J. K., & Eerola, T. (2011). The role of mood and personality in the perception of emotions represented by music. *Cortex*, *47*(9), 1099–1106.

Warrenburg, L. A. (2020). Choosing the right tune: A review of music stimuli used in emotion research. *Music Perception*, *37*(3), 240–258.

Yik, M., Mues, C., Sze, I. N. L., Kuppens, P., Tuerlinckx, F., De Roover, K., Kwok, F. H. C., Schwartz, S. H., Abu-Hilal, M., Adebayo, D. F., Aguilar, P., Al-Bahrani, M., Anderson, M. H., Andrade, L., Bratko, D., Bushina, E., Choi, J. W., Cieciuch, J., Dru, V., … Russell, J. A. (2022). On the relationship between valence and arousal in samples across the globe. *Emotion*, *23*, 332–344. https://doi.org/10.1037/emo0001095