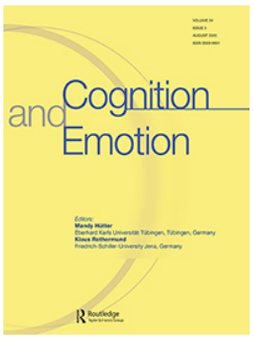

Cognition and Emotion

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/pcem20

Development of emotion recognition in popular music and vocal bursts

Dianna Vidas, Renee Calligeros, Nicole L. Nelson & Genevieve A. Dingle

To cite this article: Dianna Vidas, Renee Calligeros, Nicole L. Nelson & Genevieve A. Dingle (2020) Development of emotion recognition in popular music and vocal bursts, Cognition and Emotion, 34:5, 906-919, DOI: 10.1080/02699931.2019.1700482

To link to this article: https://doi.org/10.1080/02699931.2019.1700482

View supplementary material

Published online: 05 Dec 2019.

Submit your article to this journal

Article views: 1042

View related articles

View Crossmark data

Citing articles: 3 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=pcem20

COGNITION AND EMOTION
2020, VOL. 34, NO. 5, 906–919
https://doi.org/10.1080/02699931.2019.1700482

Routledge
Taylor & Francis Group

Check for updates

Development of emotion recognition in popular music and vocal bursts*

Dianna Vidas, Renee Calligeros, Nicole L. Nelson and Genevieve A. Dingle

School of Psychology, University of Queensland, St Lucia, Australia

ABSTRACT

Previous research on the development of emotion recognition in music has focused on classical, rather than popular music. Such research does not consider the impact of lyrics on judgements of emotion in music, impact that may differ throughout development. We had 172 children, adolescents, and adults (7- to 20-year-olds) judge emotions in popular music. In song excerpts, the melody of the music and the lyrics had either congruent valence (e.g. happy lyrics and melody), or incongruent valence (e.g. scared lyrics, happy melody). We also examined participants' judgements of vocal bursts, and whether emotion identification was linked to emotion lexicon. Recognition of emotions in congruent music increased with age. For incongruent music, age was positively associated with judging the emotion in music by the melody. For incongruent music with happy or sad lyrics, younger participants were more likely to answer with the emotion of the lyrics. For scared incongruent music, older adolescents were more likely to answer with the lyrics than older and younger participants. Age groups did not differ on their emotion lexicons, nor recognition of emotion in vocal bursts. Whether children use lyrics or melody to determine the emotion of popular music may depend on the emotion conveyed.

ARTICLE HISTORY
Received 12 December 2018
Revised 27 November 2019
Accepted 1 December 2019

KEYWORDS
Emotion recognition; development; music; lyrics; vocal bursts

The ability to recognise and integrate emotion information from a range of sources is a precursor for social competence and self-regulation (Ciarrochi, Heaven, & Supavadeeprasit, 2008). Past research has extensively examined how children acquire emotion words and concepts (Nook, Sasse, Lambert, McLaughlin, & Somerville, 2017; Russell & Widen, 2002), and develop the ability to recognise emotion in visual stimuli such as facial expressions, body posture, and pictures (Nelson & Russell, 2011). In contrast, relatively little research has focused on the development of emotion recognition in auditory stimuli, such as vocal sounds or music (Laukka & Juslin, 2007; Morton & Trehub, 2001, 2007; Vidas, Dingle, & Nelson, 2018).

Music cues to emotion

Music expresses emotion through acoustic signals that listeners perceive, recognise, and are "moved" by (Juslin & Sloboda, 2010). Participants agree when judging emotion categories in music (Coutinho & Dibben, 2013), and sensitivity to emotion-specific acoustic patterns in music emerges early in development, and improves with age (Dalla Bella, Peretz, Rousseau, & Gosselin, 2001; Dolgin & Adelson, 1990; Franco, Chew, & Swaine, 2016; Vidas et al., 2018). Children aged 3–5 years correctly identify the valence of musical melodies (i.e. happy from sad music), however are unable to distinguish emotions such as anger and fear in music (Cunningham & Sterling, 1988). Children aged 4 and 5 years can also create simple emotive songs, although use few musically expressive devices (Adachi & Trehub, 1998). At 9 years, children most accurately identify happy melodies, followed by angry and fearful melodies (Dolgin & Adelson, 1990), and happiness and sadness are easier to identify for 3- to 7-year-olds than fear and emotional neutrality (Stachó, Saarikallio, Van Zijl, Huotilainen, & Toiviainen, 2013). Even at 12 years, children

CONTACT Dianna Vidas dianna.vidas@uq.net.au
*The data that support the findings of this study are openly available on OSF at https://osf.io/kuwsj/
Supplemental data for this article can be accessed https://doi.org/10.1080/02699931.2019.1700482
© 2019 Informa UK Limited, trading as Taylor & Francis Group

more consistently label music as happy or sad, rather than using emotion labels such as exciting or calm (Kratus, 1993). Therefore, the ability to distinguish emotions in music improves with age, transitioning from valence-based to more specific emotion categories via a differentiation process, a pattern found repeatedly in previous research (Nook, Sasse, Lambert, McLaughlin, & Somerville, 2018; Widen, 2013; Widen & Russell, 2010). No research has been conducted with adolescents (13–19 years), and it remains unclear whether emotion recognition in music develops along a linear trajectory or increases markedly at a specific age.

Vocal cues to emotion

Speech prosody refers to the intonation (i.e. pitch variation) and rhythm (i.e. stress and timing) of spoken language. When communicating, speakers convey affective information both verbally, and non-verbally, where prosody can change the meaning and emotional tone of the verbal message (Aguert, Laval, Lacroix, Gil, & Bigot, 2013; Quam & Swingley, 2012). Parents tailor speech prosody to attract infants' attention and communicate positive or negative emotional information (Quam & Swingley, 2012; Scherer, 1995; Trainor, Austin, & Desjardins, 2000); although infants' sensitivity to intonation does not provide them with an adult-like understanding of speech prosody. This becomes apparent in early childhood, with children aged 3–4 years showing poor recognition of happy, sad, angry, and fearful speech prosody (Nelson & Russell, 2011), and clear improvement in accuracy between 6–7 years and 9–10 years (Spackman, Fujiki, & Brinton, 2006). At 9 years, emotion recognition of prosody is typically greater than chance, but still below adult levels (Aguert et al., 2013).

Vocal emotion can also be conveyed via vocal bursts, such as laughs, screams, sighs, and other emotional non-speech expressions (Belin, Fillion-Bilodeau, & Gosselin, 2008). Children as young as 5 years show near adult-like recognition of emotions in affect bursts (Sauter, Panattoni, & Happé, 2013), indicating that affect burst recognition precedes speech prosody, and may underlie recognition of emotion in speech and music (Vidas et al., 2018).

Acoustic cues in music and speech

In a comprehensive review, Juslin and Laukka (2003) highlighted similarities in the acoustic cues that express emotions in voice and music. Like music, happy speech is characterised by a fast rate and a high pitch, whereas sad speech is spoken with a slow rate and in a low pitch (Ilie & Thompson, 2006; Juslin & Laukka, 2003). Anger and fear, both located in the same quadrant in Russell's circumplex model (Russell, 1980), share similarities across music and speech such as fast speech rate/tempo, and high pitch level, although anger has greater variability of pitch (Juslin & Laukka, 2003). Furthermore, evidence suggests that recognition of emotions in music and speech may develop in parallel in children (Vidas et al., 2018).

Conflicting emotion cues

In both music and speech, emotion can be conveyed by the words spoken and by the melody or prosody. Often, the cues from prosody or melody and the accompanying language convey the same emotion; however, these cues may also convey conflicting emotions (e.g. saying "I'm fine" in a sad tone). When cues conflict, the listener must decide which cue is correct. When presented with conflicting prosody and speech cues (e.g. describing a happy event with sad prosody), adults rely on prosody whereas children under 9 years generally rely on the content of the words, despite having the ability to recognise prosody (Morton & Trehub, 2001).

Similarly, for music with incongruent cues, such as a happy melody with sad lyrics, adults rate the melody as more salient (Ali & Peynircioğlu, 2006). In contrast, children's tendency to rely on lyrics varies with the content. When hearing a melody sung with nonsense words, children judge the emotion based on melody. However, when the melody is accompanied by lyrics, children rely on the words (Morton & Trehub, 2007), suggesting that lyrical information dominates children's judgments.

The melody and lyrics used in previous developmental research used simple sentences and melodies (Morton & Trehub, 2007). However, the favourite music genres of young people include pop and rock, among others (Greasley & Lamont, 2011; North, Hargreaves, & O'Neill, 2000), and popular music lyrics contain poetic devices such as rhymes and repetition to communicate emotion to listeners rather than simple sentences. Thus, previous research on emotion recognition in music may not reflect how children and adolescents interact with music in their everyday lives.

Adolescents are the largest consumers of popular music (McFerran, Garrido, O'Grady, Grocke, & Sawyer,

2015; Papinczak, Dingle, Stoyanov, Hides, & Zelenko, 2015; Saarikallio & Erkkila, 2007). Music fulfils many functions for adolescents, allowing them to portray an "image" to the outside world, and satisfy their emotional needs (North et al., 2000). Young people's use of music influences their behaviours, both healthy (e.g. emotion regulation) and unhealthy (aggression, suicidality) (Andersen, Carnagey, & Eubanks, 2003; Saarikallio, Gold, & McFerran, 2015; Stack, Lester, & Rosenberg, 2012). Adolescents' heavy use of music makes it necessary to understand the relative attention paid to lyrical versus melodic cues during emotion recognition in music.

The current study

To date, few studies have examined the impacts of melody and lyrics on emotion recognition. Therefore, we conducted an exploratory study to examine how lyrics and melody in popular music influence judgements of emotion throughout childhood and adolescence. Participants were analysed by school age groups.

We used music from modern Australian music charts to capture an ecologically valid account of the cues used to infer emotion in contemporary music tracks. We also sampled a broad range of ages spanning childhood through adulthood, splitting participants into four age bands: primary school (7–11 years), middle school (12–14 years), senior school (15–17 years), and university students (18–20 years). We presented participants with music representing 4 emotions (happy, sad, angry and scared) with either congruent or incongruent affective cues. In the present study, we refer to lyrics as the content of the words of the song, with no prosodic or musical information, and the melody as the musical aspects, such as tempo and key exhibited by both the singer and the accompanying instrumentals. In the congruent trials, the lyrics of the music expressed the same emotional valence as the melody of the music (e.g. happy lyrics with a happy melody). In the incongruent trials, the 4 emotions were conveyed in the lyrics, while the melody was differently valenced (e.g. scared lyrics with a happy melody). In addition, we measured participants' ability to recognise emotion from vocal bursts expressing the same emotions.

Additional variables

Evidence shows that listener's mood influences attention to musical cues (Hunter, Schellenberg, & Griffith, 2011). Similarly, musically trained individuals have an enhanced sensitivity to the expressive cues in music (i.e. timing, pitch, and tempo) (Ilie & Thompson, 2006). Finally, the number of emotion words an individual can produce spontaneously comprises their working emotion lexicon and likely reflects their exposure to emotion words (Baron-Cohen, Golan, Wheelwright, Granader, & Hill, 2010). Although not the focus of the study, these variables were investigated to determine whether they influenced participants' emotion recognition in the musical stimuli presented. There were four hypotheses based on previous research:

H1: It was hypothesised that participants' ability to recognise emotions in congruent music and vocal stimuli would increase with age.

H2: For congruent music excerpts, it was hypothesised that there would be an effect of emotion, such that all age groups would be able to recognise happy and sad excerpts, whereas angry and scared excerpts would be less well recognised.

H3: Similarly, for incongruent music excerpts, it was hypothesised that all age groups would make lyric-consistent choices for excerpts with happy or sad lyrics more than for excerpts with angry or scared lyrics.

H4: For the incongruent music, it was hypothesised that due to children's preference for lexical content over prosody in speech, children would rely on lyrics to judge emotion, whereas adults would rely on the melody. Given the lack of evidence investigating adolescents' emotion recognition in music, we had no a priori hypothesis for this age group.

H5: It was hypothesised that there would be a correlation between the size of the emotion lexicon and use of lyrics in emotion recognition for incongruent music.

Method

Participants

In total, 172 participants aged 7–20 years (33.1% male, 66.9% female) were divided into four age groups – primary school students (N = 36; 41.2% female; 7–11 years, $M_{years} = 9.00$), middle school students (N = 26; 73.1% female; 12–14 years, $M_{years} = 13.42$), senior school students (N = 65; 81.5% female; 15–17 years, $M_{years} = 15.78$), and university students (N = 47; 61.7% female; 18–20 years, $M_{years} = 18.68$). Originally, 243 participants completed the tasks; however, 22 were removed due to incomplete data, 32 had no reported age, and 17 were outside of our target age-range.

Although a prospective power analysis was not performed, previous research had used samples of 25 participants per age group (Morton & Trehub, 2007), so we ensured that we exceeded this sample size for each of our age groups.

Materials

The study included three tasks: vocal burst recognition, musical stimuli recognition, and the emotion-lexicon task. In addition, information pertaining to participants' previous musical training, and current mood were collected.

Vocal burst stimuli

A total of eight vocal bursts – two for each emotion category of happy, sad, angry, and scared – were taken from a previously validated stimulus set (Hawk, Van Kleef, Fischer, & van Der Schalk, 2009). For each emotion category, one vocal burst was performed by a male actor and one was performed by a female actor.[1] The stimuli were sounds including laughs, grunts, and screams, with a mean duration of eight seconds.

Music stimuli

The music stimuli consisted of eight 30-second excerpts drawn from eight songs. Previous research has found that aesthetic judgements can be formed between 500–750 ms, (Belfi et al., 2018), thus 30 s was deemed sufficient time to form an emotion judgement. Table 1 outlines the excerpts used. Using the lyrics, two music excerpts were chosen by experimenters for each emotion category: happy, sad, angry, and scared. Of the two excerpts in each emotion category, one contained congruent affective cues while the other featured incongruent affective cues. For example, the happy congruent music included happy lyrics with a happy melody, while the incongruent happy music included happy lyrics with a sad melody. To ensure the music was representative of current popular Australian music[2], songs were chosen from the Triple J Countdown in 2016, and the iTunes Australia Top 100 downloaded tracks in early 2016. Most western popular music is in a major mode (Schellenberg & von Scheve, 2012), thus, six of the chosen excerpts were in a major mode, while two were in a minor mode.

Validation of melody and lyrics

Excerpts were validated by 58 adults aged 16–46, (M = 20.2), who rated the emotion conveyed in the lyrics and the melody separately. Participants were presented with 16 stimuli – 8 stimuli consisting of the song lyrics, and 8 musical excerpts. To validate the emotion in the music excerpts, karaoke versions of the stimuli (music only, no lyrics other than "oh" and "ah") were presented. To validate the emotion in the song lyrics, participants saw printed text for the same excerpts, but no music was played. All stimuli were presented for 30 s. These 16 stimuli were presented in two blocks, randomised both between and within the blocks. A confusion matrix outlining the percentage of each response for each stimulus can be found in the supplementary materials (Table S1).

For the incongruent excerpts, a majority consensus was reached by participants for 3 of the 4 songs, which also matched experimenter expectations. The angry incongruent song was removed from further analysis due to participants rating the lyrics as more scared than angry. For the congruent excerpts, the scared congruent song was removed due to no clear majority consensus on either lyrics or melody. Due to the fact that we did not look at lyrics and melody separately in any analyses of congruent stimuli, two congruent songs were retained due to reaching majority for either the lyrics or the melody – the happy congruent song (low ratings on melody, high on lyrics) and the angry congruent song (low ratings on lyrics, high ratings on melody). The inconsistency in emotion recognition between the validation test and the main study may be explained by the fact that the karaoke versions in the validation task omitted the singer's vocal prosody and presented only background melody, possibly altering participants' judgements of the emotion in the excerpts.

Additionally, the SpotifyR package (Thompson, Parry, Phipps, & Wolff, 2019) provided converging evidence supporting our categorisation of music stimuli, by analysing musical features of songs, and extracting valence and energy. The valence measure describes the musical positiveness, while the energy measure represents a measure of activity, and combines dynamic range, loudness, intensity, timbre, onset rate, and general entropy (Spotify, 2019). Table 1 shows which stimuli were included in the main analysis.

Additional variables

Musical training was measured by asking participants to respond either "yes" or "no" to the question, "have you ever had musical instrument or singing lessons". To measure current mood, participants selected a

Table 1. Table of congruent and incongruent songs.

Song Title and Artist	Year	Lyrical & Musical Emotion	BPM	Mode	Excerpt	SpotifyR – Circumplex model	Included in Main Analysis
Youth – Troye Sivan	2015	Happy lyrics, happy music	92	G major	0:21 - 0:50	Positive valence, high energy	Yes
Hello – Adele	2015	Sad lyrics, sad music	79	F minor	0:06 - 0:34	Negative valence, low energy	Yes
Paranoia – A Day to Remember	2016	Angry lyrics, angry music	150	D# minor	0:40 - 1:10	Negative valence, high energy	Yes
Emperor's New Clothes (music video) – Panic! at the Disco	2015	Scared lyrics, scared music	94	C# major	0:29 - 0:58	Positive valence, high energy	No
All of Me – John Legend	2013	Happy lyrics, sad music	63	G# major	0:52 - 1:22	Negative valence, low energy	Yes
Stitches – Shawn Mendes	2015	Sad lyrics, happy music	150	C# major	0:30 - 0:58	Positive valence, high energy	Yes
Ex's and Oh's – Elle King	2014	Angry lyrics, happy music	140	G major	2:34 - 3:03	Negative valence, high energy	No
Out of the Woods – Taylor Swift	2016	Scared lyrics, happy music	92	C major	0:51 - 1:21	Negative valence, high energy	Yes

facial expression on a five-point Likert scale, displaying faces ranging from very happy to very sad.

Procedure

The primary and middle school students were tested in a quiet area at a local museum, or in testing rooms at the university on a tablet and with headphones. Senior school students received the survey via email, and the university students completed the survey in testing rooms at the university on computers with headphones. Methods of recruitment differed due to the accessibility of different age-groups. The three tasks were presented in a set order: (1) music task (2) vocal burst task, and (3) emotion lexicon task, in order to prevent the vocal burst and the lexicon tasks from priming participants to listen to either the melody or the lyrics in music excerpts.

For the music task, the 8 music excerpts were presented with no visual information (i.e. on a blank black screen) in a randomised order. Participants were asked, "what is the mood of this song?" and selected one of the five available options of "*happy,*" "*sad,*" "*angry,*" "*scared,*" or "*none of the above.*" Participants next answered the question "is it the music or the words that tell you the mood of this song?", and selected from three possible responses: "*music,*" "*words,*" or "*none of the above.*" Participants were instructed that if they were undecided, or used both cues, they could choose which channel they relied on most when making their judgement. Adults and adolescents also rated valence, arousal, dominance, and familiarity; however, during testing, children did not have time to complete these questions, thus this was excluded before data cleaning and not analysed.

For the vocal burst task, participants listened to the eight vocal bursts, in a randomised order, and were asked "what emotion do you think this person is feeling?" Response options were "*happy,*" "*sad,*" "*angry,*" "*scared,*" or "*none of the above.*"

For the emotion lexicon task, participants were presented with an empty text box and asked to type as many emotion words as they could in 1 min. A timer embedded in the survey ensured that this limit was kept. For younger participants, responses were requested verbally and typed by the experimenter into the survey form.

Results

Main analyses

To understand the differences in responses to each emotion and type of stimulus, across age groups, we conducted a series of mixed ANOVAs. In all analyses, age group (primary school, middle school, high school, university students) was the between-groups variable, while the emotion of the stimulus was the within-groups variable. Separate ANOVAs were conducted for the 3 types of stimuli – vocal bursts, music with congruent lyrics and melody, and music with incongruent lyrics and melody. For cases in which the assumption of sphericity was violated, Greenhouse-Geisser corrections were used.

Our dependent variable was participants' recognition of the emotion presented for vocal bursts and congruent music. For the analysis of the incongruent music, because the four target emotions were conveyed by the lyrics, the dependent variable was whether or not the participant chose the emotion represented in the lyrics (1 for chose the emotion of the lyrics/"correct", 0 for did not choose the emotion of the lyrics/"incorrect"). "Accuracy" for these analyses is reported in terms of unbiased hit rate (H_u), to correct for response bias or false alarms. Unbiased hit rates are proportion scores, from 0 to 1 (where 1 is unbiased

accuracy) and were calculated using Wagner's (1993) formula. Participants' answers for congruent and incongruent music are summarised in Tables 2 and 3. The raw analyses for all stimuli are available in the supplementary materials.

Vocal Bursts

Recognition of vocal burst was above 81% for all age groups (primary school = 84%; middle school = 81.7%; senior school = 87.5%; university students = 90.8%). Contrary to our hypothesis, the 4 (age group) x 4 (emotion) ANOVA showed no main effect of age group, $F(3, 168) = .896$, $p = .445$, $\eta^2_p = .016$, and no age x emotion interaction, $F(6.22, 248.60) = 1.15$, $p = .332$, $\eta^2_p = .020$ (see Figure 1). A main effect of emotion was found, $F(2.07, 348.60) = 24.48$, $p < .001$, $\eta^2_p = .127$. Recognition was similar for happy ($M = .92$) and angry ($M = .89$) stimuli, $p = .130$, and recognition on both were higher than recognition for sad ($M = .84$) or scared ($M = .79$) stimuli, $ps \leq .001$. Scared vocal bursts were least recognised, $ps \leq .002$.

Congruent melody and lyrics

The 4 (Age Group) x 3 (Emotion: happy, sad, angry) mixed ANOVA showed that for music with congruent lyrics and melody, there was a main effect of age, $F(3, 168) = 3.59$, $p = .015$, $\eta^2_p = .060$. Least significant difference (LSD) follow up tests revealed that primary school children ($M = .57$) scored lower than middle school ($M = .72$), senior school ($M = .72$) and university ($M = .72$) students, $ps \leq .020$. All other age-groups scored similarly, $ps \geq .916$. A main effect of emotion was also found, $F(1.81, 304.12) = 31.77$, $p < .001$, $\eta^2_p = .159$. There were no differences for scores on happy ($M = .61$) and angry ($M = .57$) music, $p = .324$, while participants scored higher for sad music ($M = .87$) than both happy and angry music, $ps \leq .001$. There was no emotion by age group interaction, $F(7.52, 420.93) = 1.82$, $p = .077$, $\eta^2_p = .031$ (see Figure 2).

Table 2. Percentage of raw responses given, for music with congruent lyrics and melody.

	Emotion of song / Answered emotion	Happy	Sad	Angry	Scared
7–11yos	Happy	**64.7**	5.9	14.7	14.7
	Sad	23.5	**82.4**	0.0	2.9
	Angry	0.0	0.0	**55.9**	44.1
	Scared	2.9	5.9	11.8	**29.4**
	None of the above	8.8	5.9	17.6	8.8
12–14yos	Happy	**61.5**	0.0	0.0	3.8
	Sad	11.5	**100.0**	0.0	3.8
	Angry	0.0	0.0	**80.8**	46.2
	Scared	0.0	0.0	15.4	**19.2**
	None of the above	26.9	0.0	3.8	26.9
15–17yos	Happy	**66.2**	0.0	0.0	4.6
	Sad	15.4	**95.4**	0.0	3.1
	Angry	0.0	0.0	**89.2**	58.5
	Scared	1.5	3.1	9.2	**18.5**
	None of the above	16.9	1.5	1.5	15.4
18–20yos	Happy	**66.0**	0.0	2.2	6.4
	Sad	17.0	**100.0**	0.0	2.1
	Angry	0.0	0.0	**82.6**	40.4
	Scared	0.0	0.0	8.7	**42.6**
	None of the above	17.0	0.0	6.5	8.5

Note: Bold results indicate the expected answer for each column, where each column refers to 1 of the 4 excerpts. Due to low levels of recognition, the scared song was removed from analysis.

Incongruent melody and lyrics

The 4 (Age Group) x 3 (Emotion: happy, sad, scared) mixed ANOVA showed no main effect of age, $F(3, 168) = 2.41$, $p = .069$, $\eta^2_p = .041$, but a main effect of emotion, $F(1.78, 298.49) = 14.73$, $p < .001$, $\eta^2_p = .081$. LSD follow-up tests revealed that participants were most likely to judge the songs using the lyrics for the song with sad lyrics ($M = .41$), followed by scared lyrics ($M = .29$), then happy lyrics ($M = .19$), $ps \leq .013$.

Of particular interest to our hypotheses, the emotion by age group interaction was significant, $F(5.33, 298.49) = 6.35$, $p < .001$, $\eta^2_p = .102$(see Figure 3). As predicted, for the happy incongruent music, primary school children ($M = .31$) were more likely to use the lyrics to determine emotion than university students ($M = .14$), $p = .044$. There were no other differences between age groups, $ps \geq .073$. Similarly, for the sad incongruent music, primary school children ($M = .56$) were more likely to use the lyrics than both senior school ($M = .32$) and university students ($M = .28$), $ps \leq .002$, but scored similarly to middle school students ($M = .52$), $p = .610$. Additionally, middle school students were more likely to use the lyrics than both senior school and university students, $p \leq .028$, while the older two groups scored similarly, $p = .635$. Contrary to predictions, for the scared incongruent music, senior school students ($M = .50$) were more likely to use the lyrics to determine emotion than primary school ($M = .22$), middle school ($M = .14$), and university ($M = .30$) students, $p \leq .006$. There were no other differences between age groups, $p \geq .080$. Overall, younger children were more likely to use lyrics to judge emotion than adolescents and adults. For scared music however, adolescents were more likely to use the lyrics, suggesting a switch in which cues children use depends on the emotion conveyed.

An additional analysis was conducted coding the melody of the incongruent music as "correct".

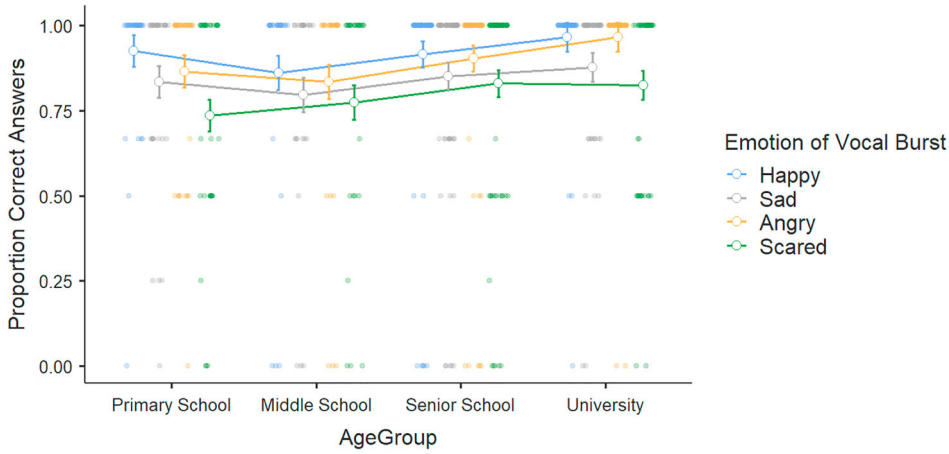

Figure 1. Average proportion of correct answers across the two vocal bursts per emotion, by age group, plus observed scores. Error bars represent standard errors.

Because three of the four incongruent stimuli had a "happy melody", the songs with negative lyrics were collapsed together into a single group, due to the use of unbiased hit-rate, which requires collapsing across stimuli of the same type. That is, for the song with happy lyrics and sad melody, we controlled just for labelling the emotion as sad, and for the other negative lyrics/happy melody stimuli, we controlled for all three responses labelling the emotion as happy.

The 4 (Age Group) x 2 (Emotion: happy lyrics/sad melody, negative lyrics/happy melody) mixed ANOVA showed no main effect of emotion, $F(1, 168) = 2.58$, $p = .110$, $\eta^2_p = .015$, but a main effect of age, $F(3, 168) = 5.49$, $p < .001$, $\eta^2_p = .089$. Follow-up tests revealed that university students ($M = .44$) were more likely to judge the songs using the melody than primary ($M = .26$), middle ($M = .22$), and senior ($M = .33$) students, $ps \leq .030$. Additionally, there was an emotion by age group interaction, $F(3, 168) = 2.84$, $p = .039$, $\eta^2_p = .048$. Reflecting results in the lyrics-consistent analysis, for the song with happy lyrics/sad melody, primary ($M = .18$) and middle ($M = .16$) students were less likely to use melody to judge emotion than senior ($M = .35$) and university ($M = .46$) students, $ps \leq .012$. For the songs with negatively valenced lyrics/happy melody, there were no significant effects. This may be due to the need to collapse across the stimuli with negative lyrics for unbiased hit rate, which may have washed out the effect. Table 3 shows how the pattern of responding

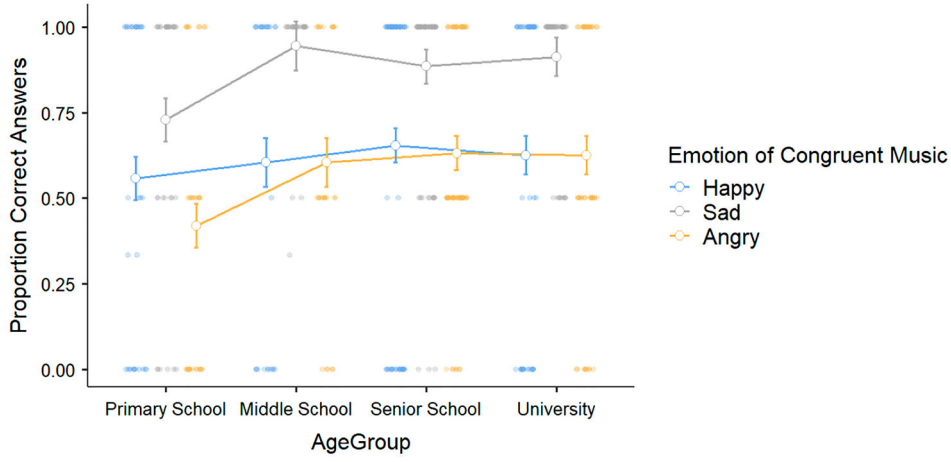

Figure 2. Proportion of correct answers for each of the three included congruent music excerpts, by age group, plus observed scores. Error bars represent standard errors.

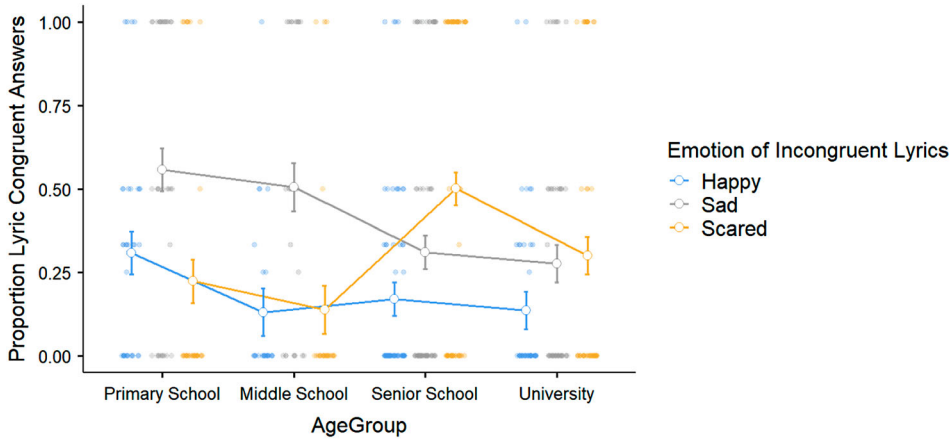

Figure 3. Average proportion of correct answers for each of the three included incongruent music excerpts, by age group, plus observed scores. Higher scores indicate using the lyrics to judge emotion. Error bars represent standard errors.

to the melody differed between the sad and scared incongruent stimuli. The results in the supplementary materials using the raw data reflect this nuanced differentiation among the songs with negatively valenced lyrics more clearly.

Emotion lexicon

Differences between age groups on the number of emotion words produced was examined in a one-way ANOVA. On average, primary school children produced 7.74 words, middle school children produced 8.50 words, senior school participants produced 8.63 words, and university students produced 9.53 words. No main effect for age groups was found, $F(3, 168) = 1.48$, $p = .222$, $\eta^2_p = .026$. Additionally, the size of emotion lexicon was correlated with age, $r = .192$, $p = .012$ but was not correlated with raw scores for recognising congruent music (scared music was excluded), $r = -.017$, $p = .826$, nor using the lyrics to determine emotion for incongruent music (angry music was excluded), $r = -.108$, $p = .159$. Table 4 shows the full correlation table.

Table 3. Percentage of raw responses given, for music with incongruent lyrics and melody.

	Emotion of lyrics / Emotion answered	Happy	Sad	Angry	Scared
7–11yos	Happy	**58.8**	*23.5*	*50.0*	*50.0*
	Sad	*26.5*	**70.6**	2.9	8.8
	Angry	0.0	0.0	**26.5**	0.0
	Scared	8.8	0.0	2.9	**23.5**
	None of the above	5.9	5.9	17.6	17.6
12–14yos	Happy	**26.9**	*19.2*	*30.8*	*42.3*
	Sad	*30.8*	**69.2**	7.7	23.1
	Angry	11.5	0.0	**23.1**	0.0
	Scared	23.1	0.0	3.8	**15.4**
	None of the above	7.7	11.5	34.6	19.2
15–17yos	Happy	**32.3**	*49.2*	*38.5*	*16.9*
	Sad	*40.0*	**36.9**	0.0	3.1
	Angry	6.2	0.0	**41.5**	6.2
	Scared	12.3	1.5	1.5	**55.4**
	None of the above	9.2	12.3	18.5	18.5
18–20yos	Happy	**27.7**	*51.1*	*36.2*	*46.8*
	Sad	*59.6*	**40.4**	4.3	2.1
	Angry	6.4	0.0	**21.3**	2.1
	Scared	2.1	0.0	14.9	**34.0**
	None of the above	4.3	8.5	23.4	14.9

Note: The emotional label for each column signifies the emotion of the music's lyrics. Bold results indicate the expected answer for the emotion of the music's lyrics, while italicised results indicate the answer matching the music's melody. Due to low levels of recognition, the song with angry lyrics was removed from analysis.

Perceived cue use

We next examined whether participants could report which emotion cue – the lyrics or the melody – they used to judge the incongruent music. We recoded participants' responses such that they scored 1 if the emotion label they selected matched the emotion of the cue they believed they used, and scored 0 if their response did not match. For example, for the song with sad lyrics and a happy melody, participants who judged the song as happy, and reported that they used the melody to make the judgment, were scored as "aligned", as were participants who responded with "sad" and "lyrics". Participants who judged the song as happy, but reported that they used the lyrics to make their judgment (or answered "sad", and "music"), were scored as "mis-aligned". Participants who had not labelled the music with either the emotion of the lyrics or the emotion of the melody, those who did

Table 4. Inter-correlations between variables.

Variables	1	2	3	4	5	6	7
1. Age							
2. Correct congruent song	.252***						
3. Correct incongruent song, melody	.216**	.105					
4. Correct incongruent song, lyrics	−.186*	−.042	−.621***				
5. Correct affect bursts	.088	.044	.167*	−.141			
6. Number of emotion words	.192*	−.017	.116	−.108	.397***		
7. Mood	−.148	−.075	.059	−.056	.206**	.168*	
8. Music Lessons	−.080	.044	.045	.067	.013	−.033	.007

Note: * $p < .05$ ** $p < .01$ *** $p < .001$. Raw scores rather than unbiased hit rate were used for correlations for ease of interpretation. For congruent music, songs were happy, sad, and angry; for incongruent music, songs had happy, sad, and scared lyrics.

not answer an emotion (none of the above), and those who answered that they used neither lyrics nor music, were not included in this analysis.

As in the main analysis, the angry incongruent song was excluded. We conducted three one-way ANOVAs, one for each incongruent music excerpt, to determine whether age group affected likelihood of correctly reporting the emotion cue they perceived they used. There were 131 participants for the happy incongruent song, 153 for the sad incongruent song, and 124 for the scared incongruent song.

For the sad incongruent music, participants in each age group were similarly likely to correctly report which cue they used, $F(3, 149) = .454$, $p = .715$, $\eta^2_p = .009$. For the scared incongruent music, participants' likelihood of correctly reporting the cue they used varied across age groups, $F(3, 120) = 7.24$, $p < .001$, $\eta^2_p = .153$. LSD follow-up tests revealed that primary students ($M = .54$) were less likely to correctly report which cue they used than middle ($M = 1.00$), senior ($M = .87$) and university students ($M = .90$), $ps < .001$. All other age groups scored similarly, $ps \geq .222$. Interestingly, the opposite pattern was found for the happy incongruent music, although all participants scored poorly in reporting which cue they used. Participants' likelihood of correctly reporting the cue they used varied across age groups, $F(3, 127) = 5.74$, $p = .001$, $\eta^2_p = .119$. LSD follow-up tests indicated that primary students ($M = .32$) were more likely to correctly report which cue they used compared with middle ($M = .00$) senior ($M = .064$) and university students ($M = .073$), $ps = .001$. All other age groups scored similarly, $ps \geq .426$. Table 5 shows the results.

Additional variables

Several variables were examined to determine their relationships with emotion recognition on our three stimuli types – vocal bursts, congruent music, and incongruent music. Because the incongruent music conveyed different emotions in the melody and in the lyrics, for these stimuli we could code responses as being correct if the song was judged in line with the emotion of the melody, or we could code responses as correct if the song was in line with the emotion of the lyrics. For these analyses, we coded the data twice, once considering the lyrics as correct, and once considering the melody as correct, and additional analyses were run with both coded versions. For this analysis, raw scores (rather than unbiased hit-rate) were used, to examine the incongruent songs easily from both perspectives. Participants' mood was examined, with those choosing "very happy" and "happy" coded as having a positive mood and given a score of 3, those choosing "neutral" given a score of 2, and those choosing "very sad" and "sad" coded as having a negative mood and given a score of 1. A more positive mood was related to generating more emotion words, $r = .17$, $p = .027$, and correctly recognising more vocal

Table 5. Percentage of responses aligned and not aligned to cue used.

		Aligned	Not aligned
Happy Incongruent	7–11 year olds	32.14	67.86
	12–14 year olds	0	100
	15–17 year olds	6.38	93.62
	18–20 year olds	7.32	92.68
Sad Incongruent	7–11 year olds	32.26	67.74
	12–14 year olds	21.74	78.26
	15–17 year olds	25	75
	18–20 year olds	20.93	79.07
Scared Incongruent	7–11 year olds	54.17	45.83
	12–14 year olds	100	0
	15–17 year olds	87.23	12.77
	18–20 year olds	89.47	10.53

Note: Aligned and not aligned columns outline the percentage of participants in each age range whose chosen emotion label for the stimulus matched or did not match the cue they claimed to use to determine emotion. Participants whose answer on emotion of song matched the cue they reported using were given a score of 1, and those whose answers did not match were given a score of 0.

burst stimuli, $r = .21$, $p = .007$. Mood was not related to emotion recognition performance in the congruent or incongruent (when lyrics or melody are correct) music tasks. Table 4 shows these correlations. Interestingly, age was associated with judging the emotion with the melody on incongruent songs, with older participants more likely to answer with the emotion of the melody, $r = .216$, $p = .004$.

Finally, 83.2% of the sample had previously had musical instrument or singing lessons. This ranged from 73.5% of the primary school group to 92.4% of the middle school group having had music lessons. Engagement in music lessons was not correlated with any focal variables.

Discussion

This study is the first to explore how children use melody and lyrics to understand the popular music they consume. We expanded upon earlier research by including additional emotions, and investigated adolescents' recognition of emotion in music. We show that earlier reports using only happy and sad music were incomplete, as emotion recognition in music differs throughout development and depends on the emotions investigated.

Vocal bursts and congruent music

As expected, participants' recognition of emotions in congruent music increased with age. Contradicting our hypothesis, however, no such increase was found for vocal bursts, corroborating research showing recognition of vocal bursts develops early in childhood (Sauter et al., 2013; Vidas et al., 2018), and suggesting that recognition of vocal bursts precedes recognition of emotion in popular music.

Incongruent music

We hypothesised that children would be more likely to use the lyrics to determine emotion in music with incongruent cues, while adults would rely on the melody. Our hypothesis was supported for happy and sad incongruent music, but not for scared music. For sad incongruent music, younger groups were more likely to use the lyrics to determine emotion than senior and university students. Similarly, for happy incongruent music, children were more likely to use the lyrics to determine emotion than university students, and the pattern of responding with the lyrics decreased with age-group. Furthermore, age was positively associated with answering with the melody on incongruent songs. These findings suggest that for happy and sad music, there is a developmental shift during adolescence towards using the melody rather than the lyrics to determine emotion, supporting Morton and Trehub's (2007) findings. As such, children's patterns of response to incongruent lyrics and melody in music is similar to their response to conflicting cues in speech (Morton & Trehub, 2001), at least for happy and sad stimuli.

Results from the scared incongruent music showed that primary, middle, and university students were less likely to answer with the lyrics than senior school students. The youngest and oldest participants likely used the emotion of the melody rather than that of the lyrics to make their judgements. It is unclear why the pattern of results for scared incongruent music differed from those of the happy and sad music. However, they reveal interesting differences between older adolescents and other age groups. Potentially, older adolescents are more drawn to the lyrics than the melody to interpret emotion from some music, which may have implications for adolescents' use of music for emotion regulation. There are several possible explanations for why the scared incongruent music did not produce the hypothesised pattern.

Firstly, it is possible that which cue individuals use develops on a U-shaped curve, rather than linearly. Research suggests that happy and sad speech prosody may be accessible in infancy but may become less salient during language acquisition in the preschool years. As children acquire language, they may redirect their limited attentional resources to the content of speech, and consequently ignore speech prosody (Gil, Hattouti, & Laval, 2016; Quam & Swingley, 2012). Eventually, categorical emotion recognition in voices improves during adolescence (Grosbras, Ross, & Belin, 2018), facilitating cross modal integration of emotion cues (Aguert et al., 2013) as well as category differentiation (Nook et al., 2018). A U-shaped curve could explain the lack of the hypothesised developmental pattern for the scared incongruent music, if the curve differs between emotion categories. For example, for happy and sad music, recognition of emotion using melody may be high in the early years, low in childhood, and high again in adolescence and adulthood. In contrast, for scared incongruent music, the curve is pushed back, such that children are less likely to use the lyrics (thus used the melody)

to determine the emotion than older adolescents, but are similar to adults.

Alternatively, the lyrics could have been interpreted differently depending on age group – the happy and sad incongruent music may have been easier for children to understand and interpret. Using popular, ecologically valid music may have been an improvement on previous research (Morton & Trehub, 2007); however, it may limited children's understanding of the content of the lyrics. If the lyrics for the scared incongruent music were too ambiguous or difficult for children to understand, they may have used the only other cue available to them to judge the emotion: the melody. This would explain why children, like adults, answered with the emotion of the melody for this song.

In contrast, older adolescents may have focused on the scared lyrics in order to process complex emotions (Dingle, Sharman, & Larwood, 2019; Papinczak et al., 2015; Sharman & Dingle, 2015; Thompson, Geeves, & Olsen, 2018). Adolescents use music for emotion identification and regulation (Dingle, Hodges, & Kunde, 2016; Saarikallio & Erkkila, 2007), thus it is crucial to understand whether the melody or the lyrics drives this process. The current results suggest that in contrast to children and adults, adolescents were drawn to the lyrics of popular music when they conveyed fear. Further research is required to understand adolescents' unique responses to these stimuli, to fully explore under what circumstances adolescents orient to the lyrics, versus the melody of the music.

Additional variables

Emotion lexicon was unrelated to recognition in congruent or incongruent music. Previous literature on emotion word comprehension has used parental report in estimating the size of children's emotion lexicon (Baron-Cohen et al., 2010); however, our methodology directly sampled children's emotion word knowledge, and represents a more ecologically valid measure of emotion lexicon. Given the mixed nature of findings related to children's emotion lexicon, additional research is needed.

Participants were not always aware of which cue they used – the lyrics or the melody – to determine the emotion of the incongruent songs. For the scared incongruent music, middle, senior, and university students were more aware than primary students of which cue they used. In contrast, for the happy incongruent music, primary students were more aware of which cue they used, although all participants scored poorly. It is unclear why this was the case for only the happy incongruent song; however, it is important to note that there are many aspects to emotion recognition in music, including evaluative conditioning, episodic memory, musical expectancy, aesthetic judgement, and more (Juslin, 2013). The inconsistency of the awareness of cue used between songs could be related to the particular songs used, the valence of the music presented, or how the lyrics and melody interacted together. Further research is needed to better understand this phenomenon.

Limitations and future research

The results of the current study replicated those found by Morton and Trehub (2007) for the emotions of happy and sad. As Morton and Trehub (2007) did not examine any other emotions, it is unclear whether our different results for scared lyrics are due to developmental factors, or the use of popular music. Regardless, the current study indicates the need for further examination of a wider range of emotion categories. Over 75% of studies using discrete emotion categories for labelling music have focused on happiness, sadness, and anger, (Eerola & Vuoskoski, 2013), ignoring emotions such as tenderness, fear, and pride, (Juslin & Laukka, 2003; Vidas et al., 2018). To gain a comprehensive understanding of how children and adolescents interpret the lyrics and melody of the music they listen to, future research should broaden the range of emotions examined, incorporating the sometimes incongruent emotions conveyed in popular music. Furthermore, the broad age range of participants required sampling from different locations (both online and in person), and resulted in uneven group numbers, thus future research is needed to test participants under similar circumstances.

Conclusions

The present study extended existing research on the development of emotion recognition in vocal bursts and music in three important ways. Firstly, it replicated research with happy and sad incongruent music, and found that the pattern varied with scared music, such that older adolescents were more likely to use the lyrics to determine emotion than children and adults. Secondly, the study included adolescents, a relatively under-researched group in musical emotion recognition, despite being the highest

consumers of popular music. This provided both corroborating and conflicting results, suggesting that replication is needed with adolescent samples. Finally, the study used popular music stimuli with lyrics, due to popular music's importance to young people (Papinczak et al., 2015), and in contrast to nearly 50% of music and emotion studies that focus on classical music (Eerola & Vuoskoski, 2013).

This study provides the first in-depth examination of how lyrics and melody in popular music influence judgements of emotion in music throughout childhood and into adolescence. We provide evidence that whether children rely on the lyrics or the melody in the music they listen to varies with age, and with the emotion of the lyrics they are listening to. The findings indicate that emotion recognition in music continues to develop through adolescence, a critical period for social development, and provide insight into the possibilities for future research on emotion recognition and regulation across development.

Notes

1. The stimuli presented were unedited and obtained directly from Hawk et al. (2009). Posers presented were: Anger: F2 and M3; Fear: F3 and M3; Joy: F1 and M4; Sadness: F3 and M3.
2. Data collection was conducted in 2016

Acknowledgements

We thank Queensland Museums for supporting our data collection, and the parents and children involved.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Dianna Vidas http://orcid.org/0000-0003-0574-5610
Nicole L. Nelson http://orcid.org/0000-0001-8299-4798
Genevieve A. Dingle http://orcid.org/0000-0002-4915-169X

References

Adachi, M., & Trehub, S. E. (1998). Children's expression of emotion in song. *Psychology of Music*, *26*(2), 133–153. doi:10.1177/0305735698262003

Aguert, M., Laval, V., Lacroix, A., Gil, S., & Bigot, L. L. (2013). Inferring emotions from speech prosody: Not so easy at age five. *PLoS ONE*, *8*(12), e83657–8. doi:10.1371/journal.pone.0083657

Ali, S. O., & Peynircioğlu, Z. F. (2006). Songs and emotions: Are lyrics and melodies equal partners? *Psychology of Music*, *34* (4), 511–534.

Andersen, C. A., Carnagey, N. L., & Eubanks, J. (2003). Exposure to violent media: The effects of songs with Violent lyrics on Aggressive Thoughts and Feelings. *Journal of Personality and Social Psychology*, *84*(5), 960–971. doi:10.1037/0022-3514.84.5.960

Baron-Cohen, S., Golan, O., Wheelwright, S., Granader, Y., & Hill, J. (2010). Emotion word comprehension from 4 to 16 years old: A developmental survey. *Frontiers in Evolutionary Neuroscience*, *2*(NOV), 1–8. doi:10.3389/fnevo.2010.00109

Belfi, A. M., Kasdan, A., Rowland, J., Vessel, E. A., Starr, G. G., & Poeppel, D. (2018). Rapid timing of musical aesthetic judgments. *Journal of Experimental Psychology: General*, *147*(10), 1531–1543. doi:10.1037/xge0000474

Belin, P., Fillion-Bilodeau, S., & Gosselin, F. (2008). The montreal affective voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research Methods*, *40*(2), 531–539. doi:10.3758/BRM.40.2.531

Ciarrochi, J., Heaven, P. C. L., & Supavadeeprasit, S. (2008). The link between emotion identification skills and socio-emotional functioning in early adolescence: A 1-year longitudinal study. *Journal of Adolescence*, *31*, 565–582. doi:10.1016/j.adolescence.2007.10.004

Coutinho, E., & Dibben, N. (2013). Psychoacoustic cues to emotion in speech prosody and music. *Cognition and Emotion*, *27*(4), 658–684. doi:10.1080/02699931.2012.732559

Cunningham, J. G., & Sterling, R. S. (1988). Developmental change in the understanding of affective meaning in music. *Motivation and Emotion*, *12*(4), 399–413. doi:10.1007/BF00992362

Dalla Bella, S., Peretz, I., Rousseau, L., & Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition*, *80*(3), B1–B10. doi:10.1016/S0010-0277(00)00136-0

Dingle, G. A., Hodges, J., & Kunde, A. (2016). Tuned in emotion regulation program using music listening: Effectiveness for adolescents in educational settings. *Frontiers in Psychology*, *7* (JUN), 1–10. doi:10.3389/fpsyg.2016.00859

Dingle, G., Sharman, L., & Larwood, J. (2019). Young people's uses of music for emotional immersion. In K. McFerran, P. Derrington, & S. Saarikallio (Eds.), *Handbook of music, adolescents, and wellbeing* (pp. 25–37).

Dolgin, K. G., & Adelson, E. H. (1990). Age changes in the ability to interpret affect in sung and instrumentally-presented melodies. *Psychology of Music*, *18*(1), 87–98.

Eerola, T., & Vuoskoski, J. K. (2013). A review of music and emotion studies: Approaches, emotion models, and stimuli. *Music Perception: An Interdisciplinary Journal*, *30*(3), 307–340. doi:10.1525/MP.2012.30.3.307

Franco, F., Chew, M., & Swaine, J. S. (2016). Preschoolers attribution of affect to music: A comparison between vocal and instrumental performance. *Psychology of Music*, *45*, 131–149.

Gil, S., Hattouti, J., & Laval, V. (2016). How children use emotional prosody: Crossmodal emotional integration? *Developmental Psychology*, *52*(7), 1064–1072. doi:10.1037/dev0000121

Greasley, A. E., & Lamont, A. (2011). Exploring engagement with music in everyday life using experience sampling methodology. *Musicae Scientiae*, *15*(1), 45–71. doi:10.1177/1029864910393417

Grosbras, M. H., Ross, P. D., & Belin, P. (2018). Categorical emotion recognition from voice improves during childhood and adolescence. *Scientific Reports*, *8*(1), 1–11. doi:10.1038/s41598-018-32868-3

Hawk, S. T., Van Kleef, G. A., Fischer, A. H., & van Der Schalk, J. (2009). "Worth a thousand words": absolute and relative decoding of nonlinguistic affect vocalizations. *Emotion*, *9*(3), 293–305. doi:10.1037/a0015178

Hunter, P. G., Schellenberg, E. G., & Griffith, A. T. (2011). Misery loves company: Mood-congruent emotional responding to music. *Emotion*, *11*(5), 1068–1072. doi:10.1037/a0023749

Ilie, G., & Thompson, W. F. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception*, *23*(4), 319–330. doi:10.1525/mp.2006.23.4.319

Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, *10*(3), 235–266. doi:10.1016/j.plrev.2013.05.008

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, *129*(5), 770–814. doi:10.1037/0033-2909.129.5.770

Juslin, P. N., & Sloboda, J. A. (2010). Introduction: Aims, organization, and terminology. In P. N. Juslin, & J. Sloboda (Eds.), *Handbook of music and emotion* (1st ed., pp. 3–12). Oxford: Oxford University Press.

Kratus, J. (1993). A developmental study of children's interpretation of emotion in music. *Psychology of Music*, *21*(1), 3–19. doi:10.1177/030573569302100101

Laukka, P., & Juslin, P. N. (2007). Similar patterns of age-related differences in emotion recognition from speech and music. *Motivation and Emotion*, *31*(3), 182–191. doi:10.1007/s11031-007-9063-z

McFerran, K. S., Garrido, S., O'Grady, L., Grocke, D., & Sawyer, S. M. (2015). Examining the relationship between self-reported mood management and music preferences of Australian teenagers. *Nordic Journal of Music Therapy*, *24*(3), 187–203. doi:10.1080/08098131.2014.908942

Morton, J. B., & Trehub, S. E. (2001). Children ' s understanding of emotion in speech. *Emotion*, *72*(3), 834–843. doi:10.1111/1467-8624.00318

Morton, J. B., & Trehub, S. E. (2007). Children's judgements of emotion in song. *Psychology of Music*, *35*, 629–639. doi:10.1177/0305735607076445

Nelson, N. L., & Russell, J. A. (2011). Preschoolers' use of dynamic facial, bodily, and vocal cues to emotion. *Journal of Experimental Child Psychology*, *110*(1), 52–61. doi:10.1016/j.jecp.2011.03.014

Nook, E. C., Sasse, S. F., Lambert, H. K., McLaughlin, K. A., & Somerville, L. H. (2017). Increasing verbal knowledge mediates development of multidimensional emotion representations. *Nature Human Behaviour*, *1*(12), 881–889. doi:10.1038/s41562-017-0238-7

Nook, E. C., Sasse, S. F., Lambert, H. K., McLaughlin, K. A., & Somerville, L. H. (2018). The nonlinear development of emotion differentiation: Granular emotional experience is low in adolescence. *Psychological Science*, *29*(8), 1346–1357. doi:10.1177/0956797618773357

North, A. C., Hargreaves, D. J., & O'Neill, S. A. (2000). The importance of music to adolescents. *British Journal of Educational Psychology*, *70*, 255–272. doi:10.1348/000709900158083

Papinczak, Z. E., Dingle, G. A., Stoyanov, S. R., Hides, L., & Zelenko, O. (2015). Young people's uses of music for well-being. *Journal of Youth Studies*, *18*(February), 1–16. doi:10.1080/13676261.2015.1020935

Quam, C., & Swingley, D. (2012). Development in children's interpretation of pitch cues to emotions. *Child Development*, *83*(1), 236–250. doi:10.1111/j.1467-8624.2011.01700.x

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. doi:10.1037/h0077714

Russell, J. A., & Widen, S. C. (2002). A label superiority effect in children's categorization of facial expressions. *Social Development*, *11*(1), 30–52. doi:10.1111/1467-9507.00185

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

Saarikallio, S., Gold, C., & McFerran, K. (2015). Development and validation of the healthy-unhealthy music scale. *Child and Adolescent Mental Health*, *20*(4), 210–217. doi:10.1111/camh.12109

Sauter, D. A., Panattoni, C., & Happé, F. (2013). Children's recognition of emotions from vocal cues. *British Journal of Developmental Psychology*, *31*(1), 97–113. doi:10.1111/j.2044-835X.2012.02081.x

Schellenberg, E. G., & von Scheve, C. (2012). Emotional cues in American popular music: Five decades of the Top 40. *Psychology of Aesthetics, Creativity, and the Arts*, *6*(3), 196–203. doi:10.1037/a0028024

Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice*, *9*(3), 235–248. doi:10.1016/S0892-1997(05)80231-0

Sharman, L., & Dingle, G. A. (2015). Extreme metal music and anger processing. *Frontiers in Human Neuroscience*, *9*(May), 272. doi:10.3389/fnhum.2015.00272

Spackman, M. P., Fujiki, M., & Brinton, B. (2006). Understanding emotions in context: The effects of language impairment on children's ability to infer emotional reactions. *International Journal of Language and Communication Disorders*, *41*(2), 173–188. doi:10.1080/13682820500224091

Spotify. (2019). Get Audio Features for a Track. Retrieved from https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/

Stachó, L., Saarikallio, S., Van Zijl, A., Huotilainen, M., & Toiviainen, P. (2013). Perception of emotional content in musical performances by 3-7-year-old children. *Musicae Scientiae*, *17*(4), 495–512. doi:10.1177/1029864913497617

Stack, S., Lester, D., & Rosenberg, J. S. (2012). Music and suicidality: A quantitative review and extension. *Suicide and Life-Threatening Behavior*, *42*(6), 654–671. doi:10.1111/j.1943-278X.2012.00120.x

Thompson, W. F., Geeves, A. M., & Olsen, K. N. (2018). Who enjoys listening to violent music and why? *Psychology of Popular Media Culture*, doi:10.1037/ppm0000184

Thompson, C., Parry, J., Phipps, D., & Wolff, T. (2019). spotifyr: R Wrapper for the "Spotify" Web API. R package version 2.1.0. Retrieved from https://cran.r-project.org/package=spotifyr.

Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science*, *11*(3), 188–195. doi:10.1111/1467-9280.00240

Vidas, D., Dingle, G. A., & Nelson, N. L. (2018). Children's recognition of emotion in music and speech. *Music & Science*, *1*, 205920431876265. doi:10.1177/2059204318762650

Wagner, H. L. (1993). On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior*, *17*(1), 3–28. doi:10.1007/BF00987006

Widen, S. C. (2013). Children's interpretation of facial expressions: The long path from valence-based to specific discrete categories. *Emotion Review*, *5*(1), 72–77. doi:10.1177/1754073912451492

Widen, S. C., & Russell, J. A. (2010). Differentiation in preschooler's categories of emotion. *Emotion*, *10*(5), 651–661. doi:10.1037/a0019005