-- Accepted for publication in *Psychology of Music* on 23/04/2016 --

PRESCHOOLERS' ATTRIBUTION OF AFFECT TO MUSIC: A COMPARISON BETWEEN VOCAL AND INSTRUMENTAL PERFORMANCE

Fabia Franco[1], Marcia Chew[2] & Joel S. Swaine[3]

[1] Department of Psychology, School of Science and Technology, Middlesex University

[2] School of Psychological Sciences, University of Melbourne

[3] School of Drama, Music and Screen, University of Hull

Corresponding author:

Fabia Franco, PhD
Department of Psychology,
School of Science and Technology,
Middlesex University,
The Burroughs, Hendon
London NW4 4BT
UK
E-mail: f.franco@mdx.ac.uk

ABSTRACT

Research has shown inconsistent results concerning the ability of young children to identify musical emotion. This study explores the influence of the type of musical performance (vocal *vs.* instrumental) on children's affect identification. Using an independent-group design, novel child-directed music was presented in three conditions: instrumental, vocal-only, and song (instrumental plus vocals) to 3- to 6-year-olds previously screened for language development (n = 76). A forced-choice task was used in which children chose a face expressing the emotion matching each musical track. All performance conditions comprised 'happy' (major mode/fast tempo) and 'sad' (minor mode/slow tempo) tracks. Nonsense syllables rather than words were used in the vocals in order to avoid the influence of lyrics on children's decisions. The results showed that even the younger children were able to identify correctly the intended emotion in music, although 'happy' music was more readily recognized and recognition appeared facilitated in the instrumental condition. Performance condition interacted with gender.

KEYWORDS:
emotion recognition in preschoolers; instrumental music; vocal music; song

When and how do we learn to perceive the intended emotional qualities in music within a given culture? In Western culture, one of the most intriguing changes in children's perception of emotion in music is the gradual development from the ages of 3 to 7 of the major mode/fast tempo association with 'happy' music, and minor mode/slow tempo connection with 'sad' music (Gabrielsson & Juslin, 2003; Gagnon & Peretz, 2003; Hunter & Schellenberg, 2010). Although mode and tempo by no means exhaust the range of expressive cues used in music to characterise different emotional qualities, a large amount of research has focused on the happy/sad distinction and their mode/tempo opposite polarization, due to the statistical consistency of these associations and the consistently high degree of recognition found across adult participants, tasks and environments (Balkwill & Thompson, 1999; Gabrielsson & Juslin, 1996; Gagnon & Peretz, 2003; Gerardi & Gerken, 1995; Peretz, Gagnon, & Bouchard, 1998; Vieillard et al., 2008). Nonetheless, while the emotion-related functions of music are evident from infancy, with mothers across different cultures singing to babies for mood regulation, play and communicating emotion (e.g., Shenfield, Trehub, & Nakata, 2003; Trehub & Nakata, 2002; Trehub & Trainor, 1998; Van Puyvelde et al., 2010), emotion-related responding to music differs between children and adults, and the literature on music perception has shown somewhat inconsistent results concerning the early stages of the developmental trajectory in the preschool years – this stage is the focus of the present study.

Appreciation of the affective connotations of music increases with age (Cunningham & Sterling, 1988; Dolgin & Adelson, 1990; Kratus, 1993; Terwogt & Van Grinsven, 1991), with general consensus that by the age of 6, children are able to identify emotions from music in a manner consistent with adult judgments (Giomo, 1993; Kratus, 1993; Terwogt & Van Grinsven, 1991; for atypical groups see Heaton, Allen, Williams, Cummins, & Happé, 2008). It remains unclear, however, how this ability develops in preschoolers. For example, Kastner and Crowder (1990) found that 3-year-olds were able to match musical excerpts to schematic faces (positive: happy or contented; negative: sad or angry) on the basis of major and minor modes. More recently, Stachó, Saarikallio, Van Zijl, Huotilainen, and Toiviainen (2013) used a forced-choice task to investigate the ability of 3-7-year-old children to interpret the emotional content expressed in music rated as emotionally neutral but manipulated through performance features to express happiness, sadness, anger, and fear. Although the 3-4-year-old children were able to recognise happiness and sadness above chance level, the overall recognition of angry and fearful expressions was poorer across all age groups (consistently with Zentner, Grandjean, & Scherer, 2008). Finally, in a study systematically manipulating mode and tempo variations of the same brief melodies, Dalla Bella, Peretz, Rousseau and Gosselin (2001) found that some reliable recognition of sad/happy music started to appear only from 5 (based on tempo) and 6-7 years based on both mode and tempo.

In sum, the ability of young children to use mode (major/minor) to match musical affect to drawings of emotional faces is supported by some studies (Kastner & Crowder, 1990; Nawrot, 2003) but not by other studies (Dalla Bella et al., 2001; Gregory, Worrall, & Sarge, 1996; Kastner & Crowder, 1990). With a few exceptions (e.g., Dolgin & Adelson, 1990), the literature suggests that the major-happy/minor-sad association is only developed after 4 years of age and becomes firmly established by the age of 7 or 8 (Dalla Bella et al., 2001; Kratus, 1993; Gerardi & Gerken, 1995). Tempo appears to be a musical expressive cue that is guiding preschoolers' recognition of emotion in music early (Crowder, Reznick, & Rosenkrantz, 1991; Dalla Bella et al., 2001). This may be based on an underlying structure (e.g., arousal) that is common with other communicative channels. For example, tempo is also associated with emotional dimensions in speech, with happy/excited speech being faster than sad/depressed speech

(Juslin & Lauuka, 2003; Morton & Trehub, 2001; Scherer, 1988).[1]

The uneven or poor performance observed in 3- and 4-year-olds is something of a paradox, given research showing some ability in infants to use cues employed in music to communicate emotion. For example, infants as young as 2 months old are able to discriminate consonance/dissonance, a cue utilized in the emotional characterization of music (at the sensorial, if not tonal level), and prefer consonant music (Trainor & Heinmiller, 1998; Trainor, Tsang, & Cheung, 2002; Zentner & Kagan, 1998; but see Plantinga & Trehub, 2014, for a note of caution). And infants aged 9 months discriminated happy/sad emotions in classical musical excerpts (Flom, Gentile & Pick, 2008), although they visually matched to the music the affectively concordant facial display only for 'happy' music (Nawrot, 2003). In the following section we will review the literature with respect to two main types of possible reasons for the paradox noted above.

Type of task, materials and type of responses (explicit vs. implicit)

First of all, it is possible that aspects of the task, musical material used or the measures chosen as dependent variables may be more or less sensitive in revealing young children's abilities. Most studies reviewed above utilized recognition tasks. However, using a production task, Boone and Cunningham (2001) highlighted that even 4-5-year-olds adjusted their movement patterns to match different emotions when asked to move a teddy bear to music. As there is converging evidence for pervasive movement/auditory interactions in the processing of rhythmic aspects of music (Trainor,

---

[1] Developmental studies of tempo preferences reveal that both preferred perceptual tempo and spontaneous motor tempo slow down as age increases (Drake, Jones, & Baruch, 2000; McAuley, Jones, Holub, Johnston, & Miller, 2006; Vanneste, Pouthas, & Wearden, 2001). Consistently, the ability of children to entrain to the musical beat is limited to a small window just below or above their preferred spontaneous motor tempo, which tends to be around ≥ 130-150 BPM (McAuley et al., 2006, from age 4; Provasi & Bobin-Begue, 2003, from the age of 2½-4 years; van Noorden, 2014, from age 3).

2007), production tasks appear promising. However, limitations in motor aptitude and other general factors (see below) make production tasks rather challenging for 3-year-olds. For instance, Boone and Cunningham's procedure involved a phase of significant modelling by the experimenter, which may be considered a confounding factor when analysing the children's responses.

Mote (2011) and Schellenberg, Nakata, Hunter, and Tamoto (2007) found superior performance in preschool children's ability to recognise emotion when using music derived from culturally familiar children's songs rather than classical or adult-directed music, thus highlighting an important influence of materials.

When lyrics and musical cues were manipulated to be consistent or inconsistent (e.g., sad story with major mode/fast tempo), 5-10-year-olds based their judgements on the lyrics rather than the musical cues in the inconsistent conditions (Morton & Trehub, 2007). This suggests a difficulty with separating language and music contents in songs, with higher attention paid to the semantic aspect. However, when Ziv and Goshen (2006) played a fast-major melody, a slow-minor melody, or no music, while 5-6-year-olds listened to an emotionally neutral story, the researchers found that children's interpretation of the emotional tone of the story was consistent with the emotion expressed by the music. It is thus possible that the lack of differentiation found by Morton and Trehub (2007) was determined by the interweaving of speech and music that characterises songs. When music was used as background to speech, children showed implicit effects of musical emotion (Ziv & Goshen, 2006). Even in these more implicit studies, though, three-year olds were not included and 4-year olds only sometimes.

Overall, these findings suggest that the type of task and materials used to test recognition of musical emotions is an important variable in young children's performance.

General developmental factors

The second type of influence having a bearing on young children's performance concerns general developmental factors, particularly when these are not controlled for in experiments. First of all, the limitations in language abilities need to be overcome when testing very young children. Children's language development contributes not only to their ability to comprehend task instructions, but also to their understanding the concept of different emotion expressions. Spackman, Fujiki, Brinton, Nelson, and Allen (2005) compared emotion understanding in children with language impairments and typically developing by using tasks requiring minimal verbal communication. Whilst no significant differences were found between the groups in recognising facial expressions (happiness, sadness, anger and fear), differences emerged in the recognition of emotion in music. Therefore, it is vital that the experimental design used with very young children does not heavily rely on verbal interpretation of emotional expressions and musical stimuli, and that language ability is controlled for.

A second general factor in music tasks is working memory, and auditory memory in particular (Vuontela et al., 2003). Working memory capacity develops significantly throughout childhood (Cowan et al., 2003; Gaillard, Barrouillet, Jarrold, & Camos, 2011; Gathercole, Pickering, Ambridge, & Wearing, 2004), hence this factor must be taken into account in the design of less demanding tasks, for instance by allowing longer exposure to stimuli.

Young children's preferences and understanding of emotional expressions in voices and faces is a third general factor that may affect performance in musical emotion recognition tasks. Despite cross-cultural variations, the existence of universals of emotional expression is largely supported for both facial and vocal affect (Byrant and Barrett, 2008; Ekman, 1972; Elfenbein & Ambady, 2002; Sauter, Eisner, Ekman, & Scott, 2010). When considering facial displays of emotion, infants show appreciation of basic emotions by 12 months of age (Camras et al., 1994; Izard et al., 1995; Walker-Andrews, 2008), with the ability to use facial expressions to make social judgements about others' emotional states improving as age increases (Cunningham & Odom, 1986; Gross & Ballif, 1991; Hosie, Gray, Russell, Scott, & Hunter, 1998; Manstead, 1993; Reichenbach & Masters, 1983). Egan, Brown, Goonan, Goonan, and Celano (1998) showed that verbal intelligence was a significant predictor in decoding facial expression of emotions and in understanding emotions through multiple modalities such as verbal content and prosody, as well as facial expression in school-age children. Most studies concur that, of the six basic emotions (happy, sad, angry, fear, surprise and disgust), a happy facial expression is the easiest to identify (Bullock & Russell, 1986; Custrini & Feldman, 1989; Grossman, Klin, Carter, & Volkmar, 2000; Phillippot & Feldman, 1990).

When considering sound, some emotion recognition abilities are clearly present in the first year of life, with even newborns discriminating positive/negative sounds and displaying a behavioural bias towards stimuli with positive valence (e.g., Mastropieri & Turkewitz, 1999). However, a recent neuroimaging study found a stronger activation in the insula and gyrus rectus when 3-7-month-olds were listening to sad vocalisations in comparison to neutral vocalisations, but similar, weaker brain activation with happy and neutral vocalisations (Blasi et al., 2011) – a process possibly associated with detecting potentially negative environmental states. Some cues of emotional identification in music are shared with cues of vocal expression of emotion, such as prosody, tempo/speech rate, loudness and timbre (e.g., Scherer, 1985, 1986). The ability to interpret these cues improves with age; for example, Morton and Trehub (2001) showed that adults rely on affective prosodic tone whereas 4-year-olds judge emotion from literal language content.

The Present Study

In sum, with respect to the gradual development from the ages of 3 to 7 of the major mode/fast tempo and minor mode/slow tempo associations with, respectively, 'happy' and 'sad' music, which was highlighted at the outset of this paper, there are methodological factors that need to be scrutinised, such as the musical features manipulated, the type of task and observed responses, as well as the type of musical materials used in experiments, in order to understand the developmental pattern. The present study aimed to address some of these concerns by (1) focusing on a distinction largely under-investigated so far, namely the comparison between instrumental music and songs; (2) using music designed to be child-directed, entirely novel for all participants, and longer in duration than in previous studies; and (3) investigating the developmental transition starting from the youngest age group credited with verbal intelligence (i.e., from the age of 3 years).

Almost invariably, studies in this field have used instrumental music from the Western classical music repertoire (see Koelsch, 2010, for a review). These pieces may be highly unfamiliar and sophisticated for young children. Studies that have not used music from this genre have still used pieces that may be unfamiliar in style to preschool-aged children, including folk songs (Kastner & Crowder, 1990), commercial jingles (Doherty, Fitzsimons, Asenbauer, & Staunton, 1999), and unfamiliar nursery songs (Gregory et al., 1996). Whilst adult-directed classical music may fail to engage children through listening during experimental procedures, it has been suggested that using music that is simpler and more familiar in style may aid to support preschoolers accuracy in musical emotion recognition (Dalla Bella et al., 2001; Mote, 2011; Schellenberg et al., 2007).

Children's first-person experience with musical instruments between 3-5 years is on average very limited, whereas children's experience with the human voice (their own and others') is extensive. Vocal play is common in infants from the babbling stage, and toddlers invent 'songs' even before they can repeat conventional songs (Davidson, 1985; Trehub, 2003). Furthermore, songs for children, similarly to infant-directed speech, show specific prosodic and phonological deviations from their adult-directed counterparts and are likely to convey emotional meaning in a more transparent way (Falk, 2011a, 2011b; Fernald & Kuhl, 1987; Kitamura & Burnham, 1998; Trainor, Austin, & Desjardins, 2000). Finally, as sung music involves familiar motor activities, it may facilitate access to meaning via an involvement of the mirror neuron system (Molnar-Szakacs & Overy, 2006). It is thus possible that the younger children would show higher emotion recognition levels with vocal than instrumental music.

In order to control for different background experiences with music, in this study only music composed ad-hoc, thus equally novel for all participants, was used. Capitalising on previous research (Mote, 2011; Schellenberg et al., 2007), only child-directed music was used to facilitate children's engagement with the happy/sad emotional contrast. The type of performance was specifically manipulated in the experiment by presenting the same melodies in three conditions: instrumental (instrumental melody + instrumental accompaniment), song (vocal melody + instrumental accompaniment) and vocal-only (a cappella solo). The vocals in the song and vocal-only conditions were identical and based on nonsense syllables (rather than lyrics) in order to control for potential effects of lyrics (Morton & Trehub, 2007). Aiming to facilitate the youngest age group tested (3 years), a simple forced-choice task was used, requiring children to match musical emotion to one of two emotional displays. However, besides emotional faces (main experiment), a control condition employing artist's situational drawings was also used with a subsample, in order to control for materials effects.

In order to facilitate children's processing of the auditory material, which might be limited by working memory capacity (Vuontela et al., 2003), another aspect adjusted in the present study was the duration of the stimuli, with musical excerpts lasting one minute, rather than 10-20 seconds as in other studies (for example, Hunter et al., 2011; Nawrot, 2003; Stachó et al., 2013; and Mote, 2011, ranging 7-35 seconds). Also, the music for the instrumental and song conditions was intentionally more complex than musically impoverished MIDI (piano) excerpts as used in other studies, as we aimed to present the children with a musical style and richness that they would be familiar with through their experience with media. However, we hypothesised that the highest facilitation would be provided by the vocal-only context, which would be more familiar to young children through both passive (being sung to, Trehub & Trainor, 1998; Trehub, Unyk, & Trainor, 1993; hearing emotional speech, Nakata & Trehub, 2004; Sambeth, Ruohio, Alku, Fellman, & Huotilainen, 2008) and active experiences (singing and vocalising emotionally themselves, Adachi & Trehub, 2011; Malloch, 2000; Welch, 2005; Young, 2002; for a review, see Trehub & Nakata, 2002).

Study 1: Adult Validation Study

As novel materials were created for the developmental sample, a generalisation experiment with adults is also included in this study. In order to accept the developmental results as valid, it was necessary to assess if adult listeners perceived the original music as having its intended emotional quality.

Method

Design

A 2 X 3 design was used, with two contrasting emotions (happy, sad) and three performance conditions (instrumental, vocal-only, song). Dependent variables were (i) accuracy and RT in an affect identification task and (ii) intensity ratings on 1 to 7 scales on three affect dimension (sad, happy, enjoyable).

Materials

Six 1-minute tracks were originally composed for children and recorded in three versions: instrumental, vocal-only, song (instrumental+vocal). The vocals were recorded in the soprano range with a trained female singer performing in child-directed style. Nonsense syllables were used to create pseudowords. 'Happy' *vs.* 'sad' affect was selected as representing clear-cut distinctions along both dimensions of arousal (high/low) and valence (positive/negative) (n=3 each, identified as H1, H2, H3 for 'happy' tracks, and S1, S2, S3 for 'sad' tracks). Overall, the 'happy' music tracks used major keys (respectively, D, G and C major), faster tempos (BPM range = 113-134), more staccato articulations, and instruments with faster attack times and shorter sustain and release times (e.g., drums, marimba, pizzicato strings), whereas the 'sad' music tracks used minor keys (respectively, G, E and A minor), slower tempos (BPM range = 75-92), more legato articulations, and instruments with darker timbres and longer sustain and release times (e.g., electric piano, string pads, church bells). The 'sad' tracks were also more dissonant and harmonically complex (e.g., with more chord changes and chord inversions). These musical-structural choices, which were made by the composer based on his intuitions of how to convey emotions through music, are generally consistent with the literature on structural factors in the perception of emotional expression in music (Gabrielsson & Lindström, 2010).

Participants

Thirty-six adults (88% women; age M = 22.2 years, SD = 7.2 years) from a multi-cultural environment took part in the experiment, with 12 participants randomly assigned to each musical performance condition (instrumental, vocal-only and song). Participants were recruited among university students who received partial course credit at Middlesex University in London (UK). None of the participants had received any formal musical training, and none had played music in any way for more than 3 years in their lifetime.

Procedure

Testing was conducted individually in a quiet room and lasted approximately 15 minutes. Stimuli were presented on a Macintosh® PowerBook G4 laptop using SuperLab software 4.0.7. Participants entered their responses using a response box Cedrus RB-730, with a. self-administered version of the task. This included two practice trials and two phases, (i) affect identification and (ii) intensity ratings for the 6 randomised musical tracks. During the first phase (i), participants were instructed to listen to each excerpt and to indicate whether it sounded 'happy' or 'sad' by button press. For each track, the choice and reaction time (RTs, henceforth) were recorded. In the second phase (ii), after listening to the six tracks again in a new random order, participants were asked to rate how sad, happy and enjoyable each track was on three 7-point rating scales (sadness, happiness and enjoyability), with 1 = not at all and 7 = extremely.

Ethical approval for both Studies 1 and 2 was granted by the Psychology Ethics Committee of Middlesex University.

Results

In the forced-choice task (i), the adult sample displayed 100% accuracy in identifying the affect of the musical tracks across all conditions. The range of RTs suggests that affect identification was quicker with some tracks (e.g., Min = 953.42 msec. with H2 in the instrumental version) than with other ones (e.g., Max = 1,969.42 msec. with S2 in the song version)[2].

The mean RTs and scores are summarised in Table 1. A 2 (affect) x 3 (performance condition) mixed ANOVA conducted on the mean RTs averaging across the three tracks of each performance type did not yield any significant effects or interactions. However, there was a moderate non-significant effect of condition $F(2,33) = 2.741$, $p = .079$, $\eta_p^2 = .142$, associated with faster decisions being taken with instrumental performance (M = 1125.06 msec.) than song (M = 1541.78 msec.) and vocal-only performance (M = 1528.64 msec.).

2 Details concerning the individual musical tracks are available in Table 1a, online supplementary materials (see Appendix 1).

When analysing the mean intensity affect ratings for happiness, sadness and enjoyment associated with individual musical tracks in each performance condition (task ii), participants clearly attributed a range of scores to each type of track, suggesting that some tracks were more successful than others in conveying the intended affect. For example, the highest happiness rating was scored by track H2 in the instrumental version (M = 6.67) and the lowest by track H3 in the song version (M = 5.67) whereas the highest sadness rating was associated with track S1/instrumental (M = 6.42) and the lowest with track S2/song (M = 5.08). A 2 (musical emotion: happy/sad) x 2 (types of rating: happiness/sadness) x 3 (condition: instrumental, vocal-only, song) mixed ANOVA was conducted on the mean scores averaging across the three tracks of each condition. Although a significant main effect of type of rating suggested a tendency to use significantly higher scores on the happiness than on the sadness rating ($F(1, 33) = 6.533$, $p = .015$, $\eta_p^2 = .165$), there was a powerful interaction between emotion and types of rating ($F(2, 66) = 327.45$, $p < .001$, $\eta_p^2 = .908$). Happiness ratings were very high with 'happy' music (M = 6.185, SD = .654) but low with 'sad' music (M = 1.361, SD = .439), and vice versa (respectively, M = 5.972, SD = .664 and M = 1.713, SD = .569). There were no significant effects of performance condition, either as a main effect or in interactions.

In order to test if fast recognition was associated with very high scores in the correct dimension, Pearson correlations were conducted between intensity ratings and RTs. The results were weak and non-significant, with the exception of two tracks. For S2, the mean sadness-score and RT in the forced-choice task were negatively correlated ($r = -.363$, $n = 36$, $p = .03$); for H3, there was a moderate correlation between the RT and the enjoyability-score (*r* =.382, *n* = 36, *p* = .02) with the happiness-score not reaching significance (*r* = .287, *n* = 36, *p* = .09)[3].

In conclusion, the adult study validated the musical materials created for the experiment as representing the target emotions whilst suggesting a range in perceived intensity of affect across different tracks.

Table 1.
Adult mean and standard error reaction time for affect identification (H, Happy, and S, Sad), and mean and standard error intensity rating for Happy vs. Sad musical tracks on happiness (H-score), sadness (S-score) and enjoyability (E-score) in each performance condition (instrumental, vocal and song).

Music	Condition	Mean RTs (msec.) (*SE*)	H-score (*SE*)	S-score (*SE*)	E-score (*SE*)
Happy	Instrumental	1096.06 (116.16)	6.25 (0.21)	1.39 (0.19)	5.39 (0.34)
	Vocal-only	1558.25 (145.37)	6.36 (0.26)	1.30 (0.18)	4.78 (0.43)
	Song	1541.78 (348.49)	5.95 (0.28)	1.39 (0.16)	4.67 (0.48)
Sad	Instrumental	1154.06 (96.73)	1.47 (0.19)	6.03 (0.31)	3.78 (0.60)
	Vocal-only	1499.03 (121.26)	1.89 (0.26)	6.14 (0.26)	3.56 (0.53)
	Song	1509.64 (222.19)	1.78 (0.22)	5.75 (0.30)	3.78 (0.53)

Study 2: Preschoolers' Recognition of Happy/Sad Affect in Music

3 Further details of the adult results are presented in Appendix 2, online supplementary materials.

The main experiment focused on the earliest stages of the developmental trajectory at which explicit behavioural tasks may be used and made specific adjustments to accommodate young children's cognitive profile as highlighted above.

Method

Design

A 3 X 2 mixed design was used, with children randomly assigned to one of three independent performance conditions (instrumental, vocal-only or song condition), with two musical emotions (sad/happy) presented within each. Three tracks were presented for each emotion, yielding a 0-3 accuracy score as dependent measure. Age (3) and sex (2) were included as additional independent factors.

Materials

The same musical stimuli were used with the developmental sample as in Study 1 (see above for the details of the stimuli).

Participants

A total of seventy-six children (49% female) took part in the experiment, aged 3 (n=27, 36–47 months, M = 42.74, SD = 3.73), 4 (n=26, 48–54 months, M = 51.50, SD = 1.79), and 5 years (n=23, 64–77 months, M = 69.52, SD = 3.37) attending nursery/school in London (UK). Thirty participants were bilinguals (39%). Children's language development was screened using the British Picture Vocabulary Scales (BPVS II, Dunn, Dunn, Whetton, & Burley, 1997), from which verbal mental age (VMA) scores were derived to identify potential participants scoring below chronological age – none were identified.

Procedure

Children were tested individually during their normal school day in a quiet room, either alone with the experimenter or with a teacher passively present. All children were given a pretest to determine whether they understood each of the two affect terms (i.e., happy and sad), for instance by asking them to recall some event that would cause someone to experience joy and sadness. Next, they were shown two facial expressions, each depicting a person exhibiting one of the above-mentioned emotional states (NimStim validated photos of facial expressions, Tottenham et al., 2009 - see Figure 1): all children identified faces correctly as ‘happy’ or ‘sad’.

Figure 1.

Example of NimStim female faces expressing sad (left) and happy (right) emotion (Tottenham et al., 2009).

Children sat facing the computer screen, beside a female experimenter who used the computer keyboard to initiate trials on a Macintosh® PowerBook G4 laptop using SuperLab software 4.0.7 for stimuli presentation, item randomization and response recording (via experimenter's key press). After familiarization, the task consisted of listening to 6 randomized different music tracks (*n* = 3 'happy', *n* = 3 'sad'), presented within a forced-choice procedure in which the researcher asked children to identify the musical emotion by naming it or pointing to one of the emotional faces displayed on the screen ("show me the face that goes with the tune"). The musical tracks were repeated when the child did not respond to them in the initial attempt. The number of trials was based on a pilot study showing that 3-year-olds were getting distracted with more trials. Testing lasted approximately 20 minutes.

In order to rule out that results could be explained by young children's difficulties with processing faces, in a separate control session, a sub-sample of 3-year-old children (*n* = 24) repeated the task but with the visual match involving situational drawings portraying 'happy' vs. 'sad' events (rather than emotional faces), see Figure 2.

Figure 2.

Artist drawings depicting sad and happy events used in the second condition © Liow.

Results

Affect Identification Accuracy

First, binomial tests were conducted on all tracks for affect identification accuracy (see Table 2), showing that children's performance was significantly above chance for the age groups of 4 and 5-6 years. Three-year-olds performed above chance only for the 'happy' tracks.

Table 2.

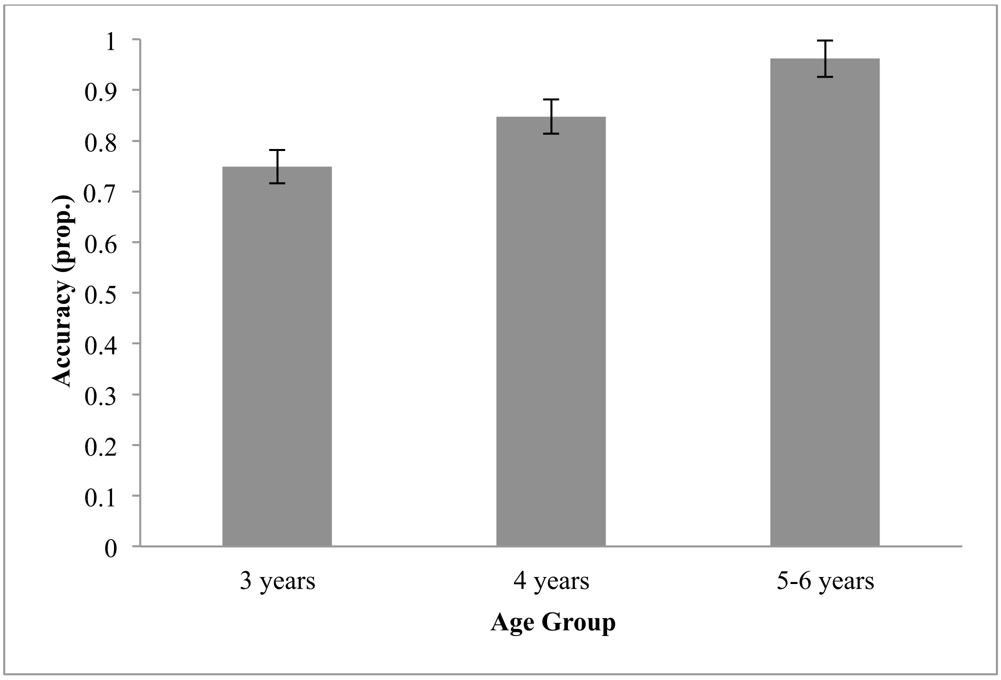
Children's results of binomial tests performed on affect identification rates (H, Happy, and S, Sad) for all music tracks.

	'Happy' tracks			'Sad' tracks		
Age Group	Track H1	Track H2	Track H3	Track S1	Track S2	Track S3
3 years (n = 27)	$p < .02$	$p < .002$	$p < .001$	NS	NS	NS
4 years (n = 26)	$p < .001$	$p < .001$	$p < .001$	$p < .002$	$p < .002$	$p < .01$
5-6 years (n = 23)	$p < .001$	$p < .001$	$p < .001$	$p < .001$	$p < .001$	$p < .001$

In order to test the effects of the design variables, the proportion of children's choices was analysed using a mixed factorial 3 x 2 x 2 x 2 GLM analysis of variance (ANOVA) with the between-subject factors of age group (3, 4, and 5 years), condition (instrumental, vocal-only and song) and gender, and repeated measures on musical affect (happy/sad). A large, significant main effect of emotion emerged, $F(1,58) = 14.755$, $p < .001$, $\eta_p^2 = .203$, with superior performance on the 'happy' tracks ($M = .909$, $SD = .199$) compared to the 'sad' tracks ($M = .797$, $SD = .258$). There was also a significant main effect of age, $F(2,58) = 10.107$, $p < .001$, $\eta_p^2 = .258$, revealing that affect identification accuracy increased with age in the expected direction (see Figure 3). Post-hoc comparisons showed that the difference in musical affect identification fell short of significance between 3-year-olds and 4-year-olds ($p = .06$), but the 5-year-olds significantly outperformed the younger age groups ($p < .001$). There were no significant effects of condition, or age x condition interaction.

Figure 3.

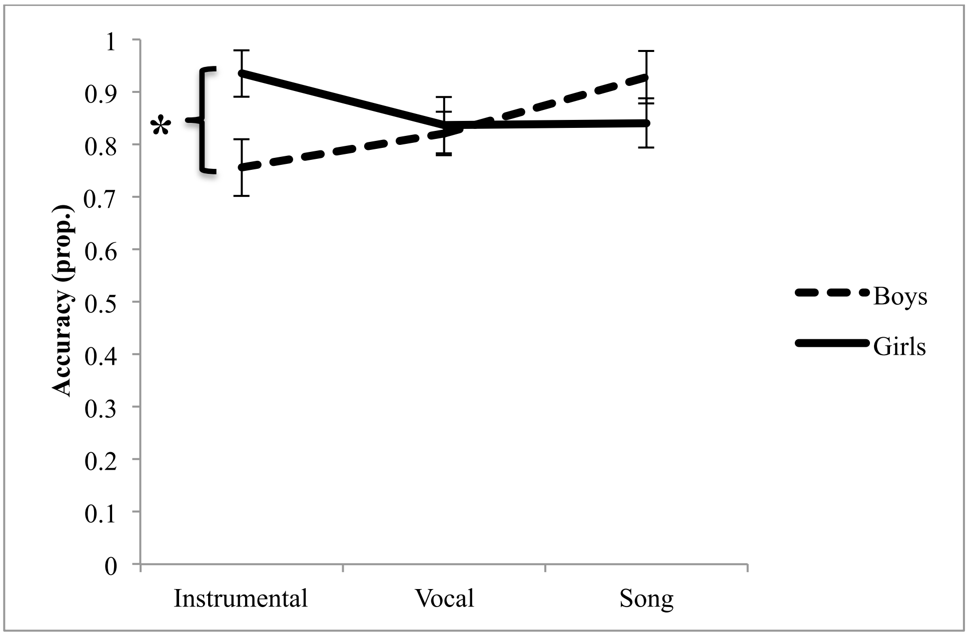
Mean accuracy scores (proportion) of the affect identification task for all age groups. Error bars represent standard error.



Gender interacted with condition, $F(2,58) = 4.007$, $p = .02$, $\eta_p^2 = .121$ - see Figure 4, with girls performing significantly more accurately than boys in the instrumental condition ($p = .017$). Boys performed significantly better in the song condition than the instrumental condition ($p = .008$), whereas girls performed significantly better in the instrumental condition than the vocal-only condition ($p = .034$).

Figure 4.

Mean accuracy scores (proportion) of affect identification for boys and girls across different conditions (instrumental, vocal-only and song). Error bars represent standard error.



Finally, in order to examine the relationship between children's verbal intelligence and the accuracy of affect identification in music, Pearson correlation coefficients between children's BPVS-derived mental age and the accuracy scores were computed. The analyses showed that verbal age was positively correlated with the overall accuracy scores in identifying affect in music, $r(76) = .328$, $p < .01$; in other words, the more advanced the children's verbal abilities were, the higher their musical emotion recognition scores were.

Control Task (Situational Drawings)

The results of the binomial tests conducted on accuracy ratings in the control condition using situational drawings instead of emotional faces are displayed in Table 3. In this case, the sub-sample of the 3-year-olds performed above chance not only for the 'happy' tracks but also for one 'sad' track.

Table 3.

Children's results of binomial tests performed on affect identification rates (H, Happy, and S, Sad) for all music tracks in the 'situational drawings' condition.

	'Happy' tracks			'Sad' tracks		
Age Group	Track H1	Track H2	Track H3	Track S1	Track S2	Track S3
3 years (n = 24)	$p < .005$	$p < .001$	$p < .005$	$p < .05$	NS	NS

Discussion

The experiment has been successful in adapting the forced-choice methodology used in other studies, so allowing very young children to show good performance in musical affect recognition from age 4 and, in part, even 3 years. Several methodological innovations were introduced, including the use of child-directed rather than adult-directed music (classical or other) as used in most studies in the field, with a few exceptions (Mote, 2011; Schellenberg et al., 2007). Child-directed music may arguably be emotionally more transparent, similarly to child-directed speech (Fernald, 1989; Trainor et al., 2000), thus supporting young children's performance. Although specific tests will be needed to fully support this hypothesis, certainly the adult sample did find affect identification extremely easy.

The main manipulation concerned the type of musical performance, varying between instrumental, song and a cappella solo. Although 3-year-olds' accuracy was 10% higher in the song condition than in the other two conditions, this difference was not statistically significant. Thus, contrary to the hypothesis, no overall facilitation for children's emotion recognition was found with performances involving vocals. This might be due to limitations in the sample size, but it is also possible that children's attention was captured by the 'strange language' used in the vocal conditions, thus obscuring rather than enhancing the intended affect. The music tracks were composed *ad-hoc,* using nonsense syllables in order to ensure that, differently from most existing studies, the songs were equally novel for all children, thus eliminating direct influences of family background (as there would be with classical music, for instance) or familiarity and semantic knowledge (as there would be with popular children's songs or children's animation movies; see Morton & Trehub, 2007). Thus, to further explore the hypothesis that vocal performance might facilitate young children's recognition of emotion in music, future studies may use single-vowel or monosyllabic vocalising ('a-a-a', 'la-la-la') rather than nonsense syllables, thereby preventing any attempts to figure out a 'meaning'.

Interestingly, an effect of performance type emerged in interaction with gender, with girls displaying higher affect identification accuracy than boys in the instrumental condition. Moreover, girls performed best in the instrumental condition whereas boys performed best in the song condition.. These results are consistent with findings from a study by Hunter, Schellenberg, and Stalinski (2011), using only instrumental music, in which boys performed significantly worse than girls in their younger age group (5 years). The combination of higher emotional sensitivity (McClure, 2000) and language development (Feldman et al., 2000) in young girls compared with boys may suggest that girls detected the emotional quality conveyed purely musically but were puzzled by novel speech-like sounds, while the boys benefitted from the predicted voice facilitation. These gender preferences are intriguing but need further testing to be fully elucidated.

The main results revealed that preschoolers can reliably recognise 'happy' and 'sad' emotion in music from the age of 4 years, and that even 3-year-olds succeed with 'happy' tracks. The 3-year-olds' failure to recognise the emotional quality of 'sad' tunes may be due to difficulties with facial displays of negative emotions (Denham & Couchoud, 1990; Gao & Maurer, 2009;).. Indeed the control condition used in this study, in which drawings of simple emotional scenes were used in the matching task (rather than faces), did allow 3-year-olds to reliably recognise at least one of the 'sad' tracks. Future research would benefit from exploring alternative tasks in order to conclusively rule out difficulties associated with face rather than music processing, for example using dynamic rather than static faces (Arsalidou, Morris, & Taylor, 2011; LaBar, Crupain, Voyvodic, & McCarthy, 2003; Wehrle, Kaiser, Schmidt, & Scherer, 2000). However, an interesting alternative explanation would suggest that, given young children's overall preferences for fast tempi (Baruch, Panissal-Vieu, & Drake, 2004; McAuley et al., 2006; Provasi & Bobin-Begue, 2003; van Noorden, 2014), 'sad' tracks based on slow tempi may simply not be engaging attention well or long enough in the youngest participants, leading to chance-level recognition. In order to test this possibility, future research may compare 'happy' tracks with music based on emotions with different valence but similar arousal levels, hence compatible with fast tempo (e.g., 'anger') – in which case the comparison would be simply based on emotional valence (e.g., conveyed by mode cues) rather than being confounded with arousal (e.g., conveyed by tempo cues) (see Franco, Swaine, Israni, Zaborowska, Kaloko, Kesavarajan,& Majek, 2014, for a design matching tempi/arousal levels across musical emotion contrasts). A final possibility is that there is a general bias towards the recognition of happy affect (e.g., see Calvo, Nummenmaa & Avero, 2010; Nawrot, 2003). Consistently, on average, adult participants provided higher happiness-ratings than their sadness-ratings, and happiness-ratings positively correlated with enjoyability-ratings.

Accuracy scores significantly positively correlated with the children's verbal age scores, and, unsurprisingly, the older children displayed superior performance, with their ability to accurately identify 'sad' emotion in the musical tracks, and. Future studies may clarify the contribution of language development (emotion vocabulary and general comprehension) to children's performance in this type of task. It is also likely that, independently from language, musical enculturation would contribute to the correct 'reading' of cues for the expression of emotion in music. Some musical parameters appear to have similar emotional effects across culture and age, such as slow tempo / minor mode for sad music, and fast tempo / major mode for happy music (Adachi, Trehub, & Abe, 2004). For instance, Western and Japanese listeners were able to correctly interpret joy, sadness and anger expressed by unfamiliar Indian ragas

(Balkwill and Thompson, 1999; Balkwill, Thompson, & Matsunaga, 2004), suggesting that listeners could appreciate affective qualities of unfamiliar music by attending to psychophysical cues. Perhaps even more cogently, Fritz et al. (2009) found that the Mafa people, living in a culturally isolated region of Cameroon, were able to identify happy, sad and scared/fearful Western music excerpts above chance level. However, Laukka, Eerola, Thingujam, Yamasaki, and Bellar (2013) found that listeners showed better performance in identifying affect from culturally familiar versus unfamiliar music, thus suggesting that perception of the affective content of music might depend on a combination of universal and culture-specific factors. This latter perspective is compatible with the idea that enculturation plays an important role in children's accuracy in recognising emotion in music composed within a given musical tradition.

The results from the adult sample showed unambiguous recognition of emotion in the musical tracks. Concerning performance type, although this factor did not show any effect on the adult ratings, the analysis of RTs during the forced-choice task revealed a near-significant effect, with instrumental music being processed up to 400 msec. faster than vocal-only music and song. As suggested for the developmental results, a possible explanation is that listeners were trying to figure out the meaning of the nonsense syllables used in the vocals. Interestingly, Brattico et al. (2011) also found that 'happy' music with no lyrics induced stronger positive emotions than 'happy' songs, as measured in behavioural ratings with adult participants, but the same difference was not revealed for 'sad' songs. Moreover, in fMRI analyses of participants listening to the same self-selected music, emotion-related areas of the brain were found to be more active during instrumental than vocal music, particularly in the case of 'happy' music. It was also found that 'sad' songs were associated with higher brain activations than 'sad' music without lyrics. This suggests the interesting possibility that affect recognition may be facilitated in different ways across different emotions.

However, the music excerpts used in Brattico et al.'s study were derived from entirely different music pieces, i.e., the happy instrumental and happy song excerpts were different pieces, rather than different performances of the same piece as in the present study.

The musical materials were obviously very limited (number of trials; number and type of emotions), and hence lacked sensitivity when considering adult cognitive abilities. However, the main aim of this study was to improve the predominant methodology in the developmental field, so that the youngest preschoolers could be tested at their best performance level. For this, the choice was made to run a limited number of trials, but to significantly increase the duration of the musical exposure with respect to previous studies. This offered the possibility that the task did not simply ask for an explicit judgement from the children, but that it actually induced the intended emotion, thus facilitating children's recognition performance (see Franco et al., 2014, for the effectiveness of a similar musical exposure and its effects on children's performance in a cognitive task). Future research would benefit from the systematic manipulation of duration of exposure in order to explore optimal levels in function of children's performance at different ages.

In summary, the present study has contributed to the development of an effective methodology to study children's affective interpretation of music, by using child-directed music, entirely novel pieces for all participants, longer durations than in previous studies for the musical exposure and by eliminating words from the vocals,. Specifically, it explored for the first time the possible role of performance conditions (purely instrumental, vocal only and mixed) on children's recognition of musical emotion. Although the results have not shown significant effects of this variable on the behavioural measure used, there was some indication that performance condition is an aspect worth investigating further using different measures, in particular implicit measures, as an interaction with gender was found in the developmental sample and a near-significant effect emerged in the generalisation data with adults (RTs). The study has also shown a likely advantage of situational drawings over facial displays of emotions in facilitating recognition of negative emotions (sadness, in this case) in young preschoolers. Thus, with adequate methodological adjustments such as those implemented here, future research will be able to fill the gap in the developmental trajectory of the ability to recognise affect in music from infancy to mid-childhood, when elements of enculturation into specific musical traditions are consolidated.

Acknowledgments

This research has been carried out with the support of British Academy grant n. SG101096 to the first author. Joel S. Swaine composed the music. Emily Weiss performed the vocals. Emily Weiss and Calley Steel contributed to piloting some of the procedures while they were on a research internship at Middlesex University babyLab, London UK (2010). We are grateful to Shen Yan Liow for producing the drawings. Finally, Middlesex University nursery, Bell Lane Primary, Oakthorpe Primary and Brookhill Nursery schools in London UK facilitated access to children for this study - their support for the research is gratefully acknowledged, together with the contribution of the children and their parents.

References

Adachi, M., & Trehub, S. E. (2011). Canadian and Japanese preschoolers' creation of happy and sad songs. *Psychomusicology: Music, Mind & Brain*, *21*, 69-82.

Arsalidou, M., Morris, D., & Taylor, M. J. (2011). Converging evidence for the advantage of dynamic facial expressions. *Brain Topography*, *24*, 149-163.

Balkwill, L., & Thompson, W. F. (1999). A cross-cultural investigation of the perception of emotion in music: psychophysical and cultural cues. *Music Perception*, *17*, 43-64.

Blasi, A., Mercure, E., Lloyd-Fox, S., Thomson, A., Brammer, M., Sauter, D., Deeley, Q., Barker, G. J., Renvali, V., Deoni, S., Gasston, D., Williams, S. C. R., Johnson, M. H., Simmons, A., & Murphy, D. G. M. (2011). Early specialization for voice and emotion processing in the infant brain. *Current Biology*, *21*, 1-5.

Boone, R. T., & Cunningham, J. G. (2001). Children's expression of emotional meaning in music through expressive body movement. *Journal of Nonverbal Behavior*, *18*, 37-55.

Brattico, E., Alluri, V., Bogert, B., Jacobsen, T., Vartiainen, N., Nieminen, S., & Tervaniemi, M. (2011). A functional MRI study of happy and sad emotions in music with and without lyrics. *Frontiers in Psychology*, *2*, 308. doi: 10.3389/fpsyg.2011.00308

Bullock, M., & Russell, J. A. (1986). Conceptual emotions in developmental psychology. In C. E. Izard & P. Read (Eds.), *Measurement of emotions in children* (Vol. 2, pp. 203-237). New York: Cambridge University Press.

Byrant, G. A., & Barrett, H. C. (2008). Vocal emotion recognition across disparate cultures. *Journal of Cognition and Culture*, *8*, 135-148.

Calvo, M. G., Nummenmaa, L., & Avero, P. (2010). Recognition advantage of happy faces in extrafoveal vision: Featural and affective processing. *Visual Cognition*, *18*(9), 1274-1297.

Camras, L. A., Oster, H., Campos, J., Campos, R., Ujiie, T., Miyake, K., Wang, L., & Meng, Z. (1998). Production of emotional facial expressions in European American, Japanese, and Chinese infants. *Developmental Psychology*, *34*, 616-628.

Cowan, N., Towse, J. N., Hamilton, Z., Saults, J. S., Elliott, E. M., Lacey, J. F., & Moreno, M. V. (2003). Children's working-memory processes: a response-timing analysis. *Journal of Experimental Psychology: General*, *132*, 113-132.

Crowder, R. G., Reznick, J. S., & Rosenkrantz, S. L. (1991). Perception of the major/minor distinction: V. Preferences among infants. *Bulletin of the Psychonomic Society*, *29*, 187-188.

Cunningham, J. G., & Odom, R. D. (1986). Differential salience of facial features in children's perception of affective expression. *Child Development*, *57*, 136-142.

Cunningham, J., & Sterling, R. (1988). Developmental changes in the understanding of affective meaning in music. *Motivation and Emotion*, *12*, 399-413.

Custrini, R. J., & Feldman, R. S. (1989). Children's social competence and nonverbal encoding and decoding of emotions. *Journal of Clinical Child Psychology*, *18*, 336-342.

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*, B1-B10.

Davidson, L. (1985). Tonal structures of children's early songs. *Music Perception*, *2*, 361-373.

Denham, S. A., & Couchoud, E. A. (1990). Young preschoolers' understanding of emotions. *Child Study Journal*, *20*, 171-192.

Doherty, C. P., Fitzsimons, M., Asenbauer, B., & Staunton, H. (1999). Discrimination of prosody and music by normal children. *European Journal of Neurology*, *6*, 221-226.

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*, 87–98.

Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition*, *77*, 251-288.

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *The British picture vocabulary scale.* Windsor: Nefer-Nelson.

Egan, G. J., Brown, R. T., Goonan, L., Goonan, B. T., & Celano, M. (1998). The development of decoding of emotions in children with externalizing behavioural disturbances and their normally developing peers. *Archives of Clinical Neuropsychology*, *13*, 383-396.

Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. In J. Cole (Ed.), *Nebraska symposium on motivation, 1971* (pp. 207-283). Lincoln: University of Nebraska Press.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: a meta-analysis. *Psychological Bulletin*, *128*, 208-235.

Falk, S. (2011a). Melodic versus intonational coding of communicative functions: a comparison of tonal contours in infant-directed song and speech. *Psychomusicology: Music, Mind and Brain*, *21,* 54-68.

Falk, S. (2011b). Temporal variability and stability in infant-directed sung speech: Evidence for language-specific patterns. *Language & Speech*, *54*, 167-180.

Feldman, H. M., Dollaghan, C. A., Campbell, T. F., Kurs-Lasky, M., Janosky, J. E., & Paradise, J. L. (2000). Measurement properties of the MacArthur Communicative Development Inventories at ages one and two years. *Child Development*, *71*, 310-322.

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development*, 1497-1510.

Fernald, A., & Kuhl, P. (1987). Acoustics determinants of infant preference for motherese speech. *Infant Behavior and Development*, *10*, 279-293.

Flom, R., Gentile, D. A., & Pick, A. D. (2008). Infants' discrimination of happy and sad music. *Infant Behavior & Development*, *31*, 716-728.

Franco, F., Swaine, J. S., Israni, S., Zaborowska, K. A., Kaloko, F., Kesavarajan, I., & Majek, J. A. (2014). Affect-matching music improves cognitive performance in adults and young children for both positive and negative emotions. *Psychology of Music*, *42*, 869-887.

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*, 68-91.

Gabrielsson, A., & Juslin, P. N. (2003). Emotional expression in music. In R. J. Davidson, K. R. Scherer & H. H. Goldsmith (Eds.), *Handbook of affective sciences*, (pp. 503-534). New York, NY: Oxford University Press.

Gabrielsson, A., & Lindström, E. (2010). The role of structure in the musical expression of emotions. In P. N. Juslin, & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications,* (pp. 367-400). Oxford, England: Oxford University Press.

Gagnon, L., & Peretz, I. (2003). Mode and tempo relative contributions to "happy-sad" judgements in equitone melodies. *Cognition and Emotion*, *17*, 25-40.

Gaillard, V., Barrouillet, P., Jarrold, C., & Camos, V. (2011). Developmental differences in working memory: where do they come from? *Journal of Experimental Child Psychology*, *110*, 469-479.

Gao, X., & Maurer, D. (2009). Influence of intensity on children's sensitivity to happy, sad, and fearful facial expressions. *Journal of Experimental Child Psychology*, *102*(4), 503-521.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Development Psychology*, *40*, 177-190.

Gerardi, G. M., & Gerken, L. (1995). The development of affective responses to modality and melodic contour. *Music Perception*, *12*, 279-290.

Giomo, C. J. (1993). An experimental study of children's sensitivity to mood in music. *Psychology of Music*, *21*, 141-162.

Gregory, A. H., Worrall, L., & Sarge, A. (1996). The development of emotional responses to music in young children. *Motivation and Emotion*, *20*, 341-348.

Gross, A. L., & Ballif, B. (1991). Children's understanding of emotion from facial expressions and situation: a review. *Developmental Review*, *11*, 368-398.

Grossman, J. B., Klin, A., Carter, A. S., & Volkmar, F. R. (2000). Verbal bias in recognition of facial emotions in children with Asperger syndrome. *Journal of Child Psychology and Psychiatry*, *41*, 369-379.

Heaton, P., Allen, R., Williams, K., Cummins, O., & Happé, F. (2008). Do social and cognitive deficits curtail musical understanding? Evidence from autism and Down syndrome. *British Journal of Developmental Psychology*, *26*, 171-182.

Hosie, J. A., Gray, C. D., Russell, P. A., Scott, C., & Hunter, N. (1998). The matching of facial expressions by deaf and hearing children and their production and comprehension of emotion labels. *Motivation & Emotion*, *22*, 293-313.

Hunter, P. G., & Schellenberg, E. G. (2010). Music and emotion. In M. R. Jones, R. R. Fay, & A. N. Popper (Eds.), *Music perception* (pp. 129–164). New York: Springer.

Hunter, P. G., Schellenberg, E. G., & Schimmack, U. (2010). Feelings and perceptions of happiness and sadness induced by music: Similarities, differences, and mixed emotions. *Psychology of Aesthetics, Creativity, and the Arts*, *4*(1), 47.

Izard, C. E., Fantauzzo, C. A., Castle, J. M., Haynes, O., Rayias, M., & Putnam, P. (1995). The ontogeny and significance of infants' facial expressions in the first 9 months of life. *Developmental Psychology*, *31*, 997-1013.

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

Kastner, M. P. & Crowder, R. G. (1990). Perception of the major/minor distinction: IV. Emotional connotations in young children. *Music Perception*, *8*, 189-201.

Kitamura, G., & Burnham, D. (1998). The infant's response to maternal vocal affect. *Advances in Infancy Research, 12*, 43-77.

Koelsch, S. (2010). Towards a neural basis of music-evoked emotions. *Trends in Cognitive Sciences*, *14*, 131-137.

Kratus, J. (1993). A developmental study of children's interpretation of emotion in music. *Psychology of Music*, *21*, 3-19.

LaBar, K. S., Crupain, M. J., Voyvodic, J. T., & McCarthy, G. (2003). Dynamic perception of facial affect and identity in the human brain. *Cerebral Cortex*, *13*, 1023-1033.

Malloch, S. N. (2000). Mothers and infants and communicative musicality. *Musicae Scientiae*, *Special Issue 1999-2000*, 29-57.

Manstead, A. S. R. (1993). Children's representation of emotions. In C. Pratt & A. F. Garton (Eds.), *Systems of representation in children: Development and use* (pp. 185–210). Chichester, England: Wiley.

Mastropieri, D., & Turkewitz, G. (1999). Prenatal experience and neonatal responsiveness to vocal expressions of emotion. *Developmental Psychobiology*, *35*, 204-214.

McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M., & Miller, M. S. (2006). The time of our lives: Life span development of timing and event tracking. *Journal of Experimental Psychology: General*, *135*, 348-367.

Molnar-Szakacs, I., & Overy, K. (2006). Music and mirror neurons: from motion to 'e'motion. *Social Cognitive and Affective Neuroscience*, *1*, 235-241.

Morton, J. B., & Trehub, S. E. (2001). Children's understanding of emotion in speech. *Child Development*, *72*, 834-843.

Morton, J. B., & Trehub, S. E. (2007). Children's judgements of emotion in song. *Psychology of Music*, *35*, 629-639.

Mote, J. (2011). The effects of tempo and familiarity on children's affective interpretation of music. *Emotion*, *11*, 618-622.

Nakata, T., & Trehub, S. E. (2004). Infants' responsiveness to maternal speech and singing. *Infant Behavior and Development*, *27*, 455-464.

Nawrot, E. S. (2003). The perception of emotional expression in music: evidence from infants, children and adults. *Psychology of Music*, *31*, 75-92.

Peretz, I., Gagnon, L., Bouchard, B. (1998). Music and emotion: perceptual determinants, immediacy, and isolation after brain damage. *Cognition*, *68*, 111-141.

Phillippot, P., & Feldman, R. S. (1990). Age and social competence in preschoolers' decoding of facial expression. *British Journal of Social Psychology*, *29*, 43-54.

Plantinga, J., & Trehub, S. E. (2014). Revisiting the innate preference for consonance. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 40-49.

Provasi, J., & Bobin-Bègue, A. (2003). Spontaneous motor tempo and rhythmical synchronisation in 2½ and 4-year-old children. *International Journal of Behavioral Development*, *27*, 220-231.

Reichenbach, L. & Masters, J. C. (1983). Children's use of expressive and contextual cues in judgments of emotion. *Child Development*, *54*, 993-1004.

Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences*, *107*(6), 2408-2412.

Schellenberg, E. G. (2012). Cognitive performance after listening to music: a review of the Mozart effect. In R. MacDonald, G. Kreutz, & L. Mitchell (Eds.), *Music, health, and wellbeing* (pp. 324-338). Oxford, UK: Oxford University Press.

Schellenberg, E. G., Nakata, T., Hunter, P. G., & Tamoto, S. (2007). Exposure to music and cognitive performance: tests of children and adults. *Psychology of Music*, *31*, 5-19.

Scherer, K. R. (1985). Vocal Affect Signaling. *Advanced Study of Behaviour*, 15, 189–244.

Scherer, K. R. (1986). Vocal affect expression: a review and a model for future research. *Psychological Bulletin*, *99*, 143-165.

Scherer, K. R. (1988). Criteria for emotion-antecedent appraisal: a review. In V. Hamilton, G. H. Bower, & N. H. Frijda (Eds.), *Cognitive perspectives on emotion and motivation* (pp. 89–126). Dordrecht: Kluwer.

Sembeth, A., Ruohio, K., Alku, P., Fellman, V., & Huotilainen, M. (2008). Sleeping newborns extract prosody from continuous speech. *Clinical Neurophysiology*, *119*, 332-341.

Shenfield, T., Trehub, S. E., & Nakata, T. (2003). Maternal singing modulates infant arousal. *Psychology of Music*, *31*, 365-375.

Spackman, M. P., Fujiki, M., Brinton, B., Nelson, D., & Allen, J. (2005). The ability of children with language impairment to recognize emotion conveyed by facial expression and music. *Communication Disorders Quarterly*, *26*, 131-143.

Stachó, L., Sarrikallio, 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*, 495-512.

Terwogt, M. M., & Van Grinsven, F. (1991). Musical expression of moodstates. *Psychology of Music*, *19*, 99-109.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., Westerlund, A., Casey, B. J., & Nelson, C. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Research*, *168*, 242-249.

Trainor, L. (2007). Do preferred beat rate and entrainment to the beat have a common origin in movement? *Empirical Musicology Review*, *2*, 17-20.

Trainor, L. J., & Heinmiller, B. M. (1998). The development of evaluative responses to music: infants prefer to listen to consonance over dissonance. *Infant Behavior & Development*, *21*, 77-88.

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*, 188-195.

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*, 188-195.

Trainor, L. J., Tsang, C. D., & Cheung, V. H. W. (2002). Preference for sensory consonance in 2- and 4-month-old infants. *Music Perception*, *20*, 187-194.

Trehub, S. (2003). The developmental origins of musicality. *Nature Neuroscience*, *6*, 669-673.

Trehub, S., & Nakata, T. (2002). Emotion and music in infancy. *Musicae Scientiae*, *Special Issue 2001-2002*, 37-61.

Trehub, S. E., & Trainor, L. J. (1998). Singing to infants: lullabies and play songs. *Advances in Infancy Research*, *12*, 43-77.

Trehub, S. E., Unyk, A. M., & Trainor, L. J. (1993). Adults identify infant-directed music across culture. *Infant Behavior & Development*, *16*, 193-211.

van Noorden, L. (2014). Auto-correlation and entrainment in the synchronous reproduction of musical pulse: developments in childhood. *Procedia – Social and Behavioral Sciences*, *126*, 117-118.

Van Puyvelde, M., Vanfleteren, P., Loots, G., Deschuyffeleer, S., Vinck, B., Jacquet, W., & Verhelst, W. (2010). Tonal synchrony in mother-infant interaction based on harmonic and pentatonic series. *Infant Behavior and Development*, *33*, 387-400.

Vanneste, S., Pouthas, V., & Wearden, J. H. (2001). Temporal control of rhythmic performance: a comparison between young and old adults. *Experimental Aging Research*, *27*, 83-102.

Vieillard, S., Peretz, I., Gosselin, N., Khalfa, S., Gagnon, L., & Bouchard, B. (2008). Happy, sad, scary and peaceful musical excerpts for research on emotions. *Cognition and Emotion*, *22*, 720-752.

Vuontela, V., Steenari, M. -R., Carlson, S., Koivisto, J., Fjällberg, M., & Aronen, E. T. (2003). Audiospatial and visuospatial working memory in 6-13 year old school children. *Learning & Memory, 10*, 74-81.

Walker-Andrews, A. S. (2008). Intermodal emotional processes in infancy. In M. Lewis, J. M. Haviland-Jones, and L. F. Barrett (Eds.), *Handbook of emotions* (pp. 364-375). NY: Guilford Press.

Wehrle, T., Kaiser, S., Schmidt, S., & Scherer, K. R. (2000). Studying the dynamics of emotional expression using synthesized facial muscle movements. *Journal of Personality and Social Psychology*, *78*, 105-119.

Welch, G. F. (2005). Singing as communication. In D. Miell, R. MacDonald, & D. J. Hargreaves (eds), *Musical communication* (pp.239-259). New York: Oxford University Press.

Young, S (2002). Young children's spontaneous vocalizations in free play: Observation of two- to three year olds in a day care setting. *Bulletin of the Council for Research in Music Education*, *152*, 43-53.

Zentner, M., & Kagan, J. (1998). Infants' perception of consonance and dissonance in music. *Infant Behavior & Development, 21*, 483-492.

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: characterization, classification, and measurement. *Emotion, 8*, 494-521.

Ziv, N., & Goshen, M. (2006). The effect of 'sad' and 'happy' background music on the interpretation of a story in 5 to 6-year-old children. *British Journal of Music Education, 23*, 303-314.

APPENDIX: SUPPLEMENTARY ONLINE MATERIALS

Appendix 1 - Table 1a.

Mean reaction time for affect identification (H, Happy, and S, Sad) and mean intensity rating for individual music tracks (H, Happy, and S, Sad) on happiness (H-score), sadness (S-score) and enjoyability (E-score) ratings in each condition (instrumental, vocal and song) in the adult validation sample.

Track	Condition	Mean RTs (msec.)	H-score	S-score	E-score
H1	Instrumental	1174.17	6.17	1.67	5.42
	Vocal-only	1491.50	6.33	1.33	4.83
	Song	1083.25	6.00	1.50	4.50
H2	Instrumental	953.42	6.67	1.17	5.50
	Vocal-only	1579.25	6.58	1.08	4.33
	Song	1373.42	6.17	1.17	5.08
H3	Instrumental	1160.58	5.92	1.33	5.25
	Vocal-only	1604.00	6.17	1.50	5.17
	Song	1243.67	5.67	1.50	4.42
S1	Instrumental	1243.25	1.33	6.42	3.50
	Vocal-only	1499.25	1.83	6.33	3.58
	Song	1221.83	1.33	6.17	3.17
S2	Instrumental	1151.42	1.67	6.08	3.92

	Vocal-only	1455.00	2.08	6.08	3.67
	Song	1960.42	2.33	5.08	4.25
	Instrumental	1067.50	1.42	5.58	3.92
S3	Vocal-only	1542.83	1.75	6.00	3.42
	Song	1346.67	1.67	6.00	3.92

Appendix 2 - Further details on adult results

The enjoyability ratings were analysed separately (mixed ANOVA emotion X condition), revealing only a significant effect of emotion, $F(1, 33) = 9.8$, $p = .004$, $\eta_p^2 = .229$, associated with overall higher enjoyability scores attributed to 'happy' than 'sad' tracks (respectively, $M = 4.944$ and $M = 3.704$).

In order to assess the relationship between the affect rating scores (H/S, H/E and S/E) within each performance condition and emotion, Pearson correlation coefficients were computed for each music track. Although there was a general pattern of negative correlations between the happiness and sadness ratings, and positive correlations between happiness and enjoyability ratings, performance conditions and individual tracks appeared to be associated with more nuanced distinctions (see Table App.2).

Table App.2.

Pearson's correlation coefficients between the affect rating scores (H, happiness; S, sadness; E, enjoyability) for individual musical tracks in all conditions (instrumental, vocal-only and song)

Condition	Track	Affect ratings	*r*	*p*	*n*
Instrumental	1H	H/S	-.821	.001**	12
		H/E	.641	.025*	12
		S/E	-.369	.238	12
	2H	H/S	-.632	.027*	12
		H/E	.510	.091	12

		S/E	-.430	.163	12
	3H	H/S	-.645	.023*	12
		H/E	.528	.078	12
		S/E	-.617	.033*	12
	S1	H/S	-.645	.023*	12
		H/E	-.132	.682	12
		S/E	.244	.444	12
	S2	H/S	-.735	.006**	12
		H/E	-.127	.695	12
		S/E	.239	.455	12
	S3	H/S	.122	.705	12
		H/E	-.320	.310	12
		S/E	-.189	.557	12
Vocal-only	1H	H/S	-.681	.015*	12
		H/E	.431	.162	12
		S/E	-.293	.355	12
	2H	H/S	-.746	.005*	12
		H/E	-.029	.928	12
		S/E	-.067	.835	12
	3H	H/S	-.812	.001**	12
		H/E	.529	.077	12
		S/E	-.225	.481	12
	S1	H/S	.111	.730	12
		H/E	.621	.031*	12
		S/E	.131	.686	12

	S2	H/S	-.743	.006*	12
		H/E	.472	.121	12
		S/E	-.069	.830	12
	S3	H/S	-.686	.014*	12
		H/E	-.369	.237	12
		S/E	-.453	.139	12
Song	1H	H/S	-.849	.001**	12
		H/E	.218	.497	12
		S/E	-.103	.751	12
	2H	H/S	-.302	.339	12
		H/E	.636	.026*	12
		S/E	-.808	.001**	12
	3H	H/S	-.456	.137	12
		H/E	.472	.121	12
		S/E	-.251	.432	12
	S1	H/S	-.279	.380	12
		H/E	.327	.300	12
		S/E	-.137	.670	12
	S2	H/S	-.182	.571	12
		H/E	-.115	.772	12
		S/E	-.209	.514	12
	S3	H/S	-.245	.443	12
		H/E	-.276	.386	12
		S/E	.315	.319	12

* Correlation significant at the .05 level (2-tailed). **Correlation significant at the .01 level (2-tailed).

In the instrumental condition, all three ‘happy’ tracks presented a significant negative correlation between mean happiness and sadness ratings by the participants. Moreover, the mean happiness score was positively correlated with the mean enjoyability score, but this correlation was fully significant only for H1. Similarly, sadness scores were negatively correlated with enjoyability, but again this was significant only for H2. As for the ‘sad’ tracks, there was again a negative correlation between happiness and sadness ratings (significant for S1 and S2). In the vocal-only condition, the correlations for both the ‘happy’ (all three) and ‘sad’ tracks (S2 and S3) showed again a strong inverse correlation between the mean happiness and sadness rating scores. However, the correlations with the mean enjoyability ratings were weak and non-significant, except in S1 where there was a significant relationship between happiness and enjoyability ratings. Finally, the song condition presented fewer significant associations than the previous performance conditions, although the pattern of results was similar. Interestingly, the ‘happy’ song H2 showed positive significant correlations for both happiness/enjoyability and sadness/enjoyability ratings, although the negative relationship happiness/sadness score was not significant.