Developmental Science

Developmental Science 19:6 (2016), pp 933–946 DOI: 10.1111/desc.12348

PAPER

Emotional experience in music fosters 18-month-olds′ emotion–action understanding: a training study

Tik Sze Carrey Siu and Him Cheung

Department of Psychology, The Chinese University of Hong Kong, Hong Kong.

Abstract

We examine whether emotional experiences induced via music-making promote infants' use of emotional cues to predict others' action. Fifteen-month-olds were randomly assigned to participate in interactive emotion training either with or without musical engagement for three months. Both groups were then re-tested with two violation-of-expectation paradigms respectively assessing their sensitivity to some expressive features in music and understanding of the link between emotion and behaviour in simple action sequences. The infants who had participated in music, but not those who had not, were surprised by music–face inconsistent displays and were able to interpret an agent's action as guided by her expressed emotion. The findings suggest a privileged role of musical experience in prompting infants to form emotional representations, which support their understanding of the association between affective states and action.

Research highlights

- Three months of interactive emotion training and musical engagement beginning at 15 months of age leads to enhanced perception of musical expressiveness.
- Emotion training supported by musical activities also promotes 15- to 18-month-olds' understanding of others' affective states behind behaviour.
- Infants' sensitivity to musical expressiveness and affective interpretation of behaviour may share the same underlying mechanism.

Introduction

Music has long been regarded as a language of emotion which facilitates communication of affective information across individuals (Cooke, 1959; Sloboda & Juslin, 2001). In particular, caregivers worldwide sing and speak in an exaggerated tone to engage their preverbal infants in communication (Fernald, 1991; Trehub, 2007). Parents render distinct types of songs in different performance styles and infants respond to them differently (Rock, Trainor & Addison, 1999; Trehub & Trainor, 1998). Such 'musical' interaction sets the stage for emotion sharing which nurtures the parent–infant bond (Dissanayake, 2000; Falk, 2004; Trehub, 2001). Furthermore, music engenders affect and modulates infants' arousal (Shenfield, Trehub & Nakata, 2003). Nevertheless, it remains unknown whether the emotional experience induced by music impacts on infants' developing understanding of the social world. Do infants capitalize on this first-person emotional experience via a process of analogy to interpret others' affective states? In the present paper we report a training study in which 15-month-olds were engaged in interactive activities for emotional experience either with or without participatory music-making. After three months of training, we re-tested their sensitivities to some expressive features in music and also compared their abilities to infer emotion to interpret simple action sequences in a separate task. We asked whether engagement in music would promote infants' understanding of the link between affective states and behaviour.

Music is unique in communicating emotion. It allows performers to express emotion and evokes strong affective experience among listeners (Budd, 1985; Davies,

Address for correspondence: Him Cheung, Department of Psychology, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong; e-mail: hcheung@psy.cuhk.edu.hk

2001; Gabrielson & Juslin, 2003), and this is seen as the primary purpose of music engagement by adult listeners (Juslin & Laukka, 2004). The six basic emotions of happiness, anger, disgust, surprise, sadness, and fear are thought to be readily represented in music, although happiness and sadness are easier to be expressed and identified (Mohn, Argstatter & Wilker, 2010; Vieillard, Peretz, Gosselin, Khalfa, Gagnon *et al.*, 2008). Research has shown that emotion is coded in music by basic acoustic and musical-system-specific cues (see Trainor & Corrigall, 2010, for a review). For instance, pitch, rate, amplitude, rhythm, and timbre are basic acoustic cues conveying meaning in both music and speech (e.g. Juslin & Laukka, 2003; Bhatara, Tirovolas, Duan, Levy & Levitin, 2011), whereas specific interval relationships defining the major and minor mode are pitch-based cues specific to the Western tonal system (e.g. Gabrielsson & Juslin, 2003). Prior studies with adults and older children have indicated that the intended emotions in music are understood by the listeners. Adults across cultures and musical backgrounds are accurate at identifying basic emotions expressed in both vocal and instrumental music (Balkwill & Thompson, 1999; Balkwill, Thompson & Matsunaga, 2004; Bigand & Poulin-Charronnat, 2006; Gagnon & Peretz, 2003; Husain, Thompson & Schellenberg, 2002; Juslin & Laukka, 2003). Children as young as 4 to 5 years old are able to pick the correct schematic face that depicts the same emotion behind the music heard (Dalla Bella, Peretz, Rousseau & Gosselin, 2001; Dolgin & Adelson, 1990; Kratus, 1993; Mote, 2011). Although young infants do not show a consistent preference for emotionally concordant music–face displays (Nawrot, 2003), Flom and colleagues have demonstrated that infants at 5 months could reliably discriminate happy and sad musical excerpts (Flom, Gentile & Pick, 2008; Flom & Pick, 2012). It has also been shown that 6- to 7-month-olds prefer high-pitched playsongs with a fast tempo and a low-pitched version of lullabies (Conrad, Walsh, Allen & Tsang, 2011; Tsang & Conrad, 2010; Volkova, Trehub & Schellenberg, 2006). This context-dependent preference for pitch and tempo indicates infants' sensitivity to some of the contrastive cues used in music to communicate emotions.

Although infants may be sensitive to acoustic cues representing emotion in music, whether music evokes affective experience is another issue. We think that infants may indeed feel the expressed emotion in music. It was shown that 6-month-olds directed their attention inward during lullabies but exhibited more outward behaviours to the external environment during playsongs, suggesting some sensitivity to the emotional content of music (Rock *et al.*, 1999). Zentner and Kagan (1998) also found more vocalizing and less fretting among 6-month-olds who listened to consonant melodies having a positive affective connotation. More importantly, infant-directed singing regulated infants' emotional arousal as indexed by salivary cortisol level (Shenfield *et al.*, 2003). The physiological and behavioural changes following exposure to music hence suggest that music can induce active emotional experience in infants.

But how do infants first get to understand emotions expressed elsewhere? A number of studies have demonstrated that infants' hands-on experience with the world promotes their understanding of similar mind states and behaviours in others (e.g. Brune & Woodward, 2007; Cicchino & Rakison, 2008; Johnson, Davidow, Hall-Haro & Frank, 2008; Loucks & Sommerville, 2012; Needham, 2000; Perone, Madole, Ross-Sheehy, Carey & Oakes, 2008; Sommerville & Woodward, 2005; Woodward & Guajardo, 2002). Sommerville, Woodward and Needham (2005) outfitted pre-reaching 3-month-olds with Velcro-covered sticky mittens to facilitate their apprehension of objects in a training session. The infants were subsequently shown an agent grasping some objects with a sticky mitten in a visual habituation paradigm (Woodward, 1998). The infants who had explored the objects with sticky mittens appraised the action as goal-directed by showing a stronger novelty response to the agent grasping a new object. Sommerville, Hildebrand and Crane (2008) argued that it is the first-person experience that contributes to such action understanding, and the facilitatory effect could not be achieved via matched observational experience. Some recent work further found that this mitten experience also accelerated 4.5-month-olds' understanding of causality (Rakinson & Krogh, 2012) and 3-month-olds' spontaneous orienting towards objects and faces (Libertus & Needham, 2011), suggesting that action experience exerts a causal impact on understanding the physical and social world.

Some other research has investigated how changes in visual experience alter infants' understanding of others' perceptual status in goal-directed actions. Meltzoff and Brooks (2008) let 12- and 18-month-olds wear an ordinary blindfold so that they could not see, and also had control infants see and feel a blindfold without using it or wear a 'trick' blindfold through which they could see. After these experiences, all the infants were shown a blindfolded adult agent facing an object. The infants who had experienced how an ordinary blindfold worked followed the agent's gaze less than the control infants whose view had not been obstructed. Hence the blindfolded infants seemed to know better that the blindfolded agent's 'gaze' did not have anything to do with actual seeing. Myowa-Yamakoshi, Kawakita, Okanda and Takeshita (2011) demonstrated that 12-month-olds who had been blindfolded looked longer at a blindfolded agent performing successful than failed goal-directed actions, supposedly because the infants' own blindfolding experience made the blindfolded agent's successful attempts particularly surprising and thus interesting to look at. Other studies using the eye-tracking technique have established a direct link between infants' own action experience and anticipation of others' action (Cannon & Woodward, 2012; Cannon, Woodward, Gredeback, von Hofsten & Turek, 2012; Gredeback & Melinder, 2010). Overall, the evidence points to a privileged role of self-experience, motor or visual in nature, in infants' interpretation of others' goal-directed actions.

Meltzoff (2007) puts forward the 'like me' hypothesis stipulating that infants are able to map from self to other and vice versa. The recognition of self–other equivalences enables infants to reflect on their own internal states as if the action were self-produced when they see others perform the same action. In other words, infants' own experience with the action supplements their perception of others' similar action. In the same vein, we argue that an emotional experience elicited via music-making is likely to provide additional proprioceptive information for understanding others' emotion-guided action. To date, it remains an open question as to the age at which infants start to relate emotional information to others' behaviour (Barna & Legerstee, 2005; Phillips, Wellman & Spelke, 2002; Repacholi & Gopnik, 1997; Vaish & Woodward, 2010). Phillips *et al*. (2002) demonstrated that 12-month-olds were able to anticipate an agent's subsequent grasping of an object which (s)he had looked at and shown some positive emotion about. Vaish and Woodward (2010) argued that such anticipation could be due to the mere fact that the agent had gazed at the target object, not requiring any emotion understanding. Vaish and Woodward (2010) thus used both happiness and disgust in their study to assess infants' ability to link emotion to action. Regardless of the emotion expressed, 14-month-olds demonstrated stronger novelty responses when the actor reached for the previously unattended cup. The results suggest that infants at this age refer to eye-gaze, rather than emotional display, in anticipating others' action. Nevertheless, Hepach and Westermann (2013) recently found an increase in 14-month-olds' sympathetic arousal, measured by pupil dilation, in response to incongruent emotion–action events, indicating their sensitivity to emotion–action incongruence and an ability to interpret actions based on facial emotional expressions.

The above studies, however, did not ask what kind of experience might facilitate such emotion–action understanding. Given the growing evidence that infants' own experience impacts on their construal of others' action, we hypothesize that first-person exploration of emotion in music facilitates infants' linking of behavioural cues to emotional states in interpreting others' action. In a recent study, Gerry, Unrau and Trainor (2012) showed that infants in an active music-making class, compared to those listening to music passively, showed more smiling and laughter, were less distressed to limitations and novel situations, and were easier to be soothed after the music training. Kirschner and Tomasello (2010) further demonstrated in 4-year-olds that making music in a group evidenced more prosocial behaviour including spontaneous helping and cooperative problem-solving. Similarly, Cirelli, Einarson and Trainor (2014) found that 14-month-olds who were held and bounced to music in synchrony with a bouncing adult were more likely to later display altruistic behaviour towards that adult than those bounced out of synchrony with the adult. The present study extends the effect of active musical experience from promoting actual social skills and other specific behaviour indicative of better social-emotional functioning to enhancing more abstract understanding of the emotion–action link. We predict an early effect of music-making in relation to emotion learning on infants' general understanding of how emotion and behaviour are related.

To examine this possibility, we randomly assigned 15-month-olds to either a music or control training condition and looked at the impact of music engagement on their sensitivity to musical expression and ability to predict an agent's action based on her emotional display. Training in both groups was structured to introduce facial, bodily, and vocal expressions that we typically use to communicate basic emotions. While learning about how emotions were expressed, the infants and parents in the music group also participated in interactive, supportive musical activities such as singing, playing toy percussion instruments, and dancing to a collection of action songs and lullabies. This training procedure was modelled on the Suzuki Early Childhood Education approach which highlights the involvement of parents and peers in children's music-making (Jones, 2004). In contrast, the infant-parent dyads in the control group received similar training on emotion communication but with no music. Because music engenders strong affective experience and constitutes an integral part of infants' early communication, we think that engagement in music functions as an effective representation medium of emotion that infants can draw on to understand others' emotion-directed actions. Hence, we hypothesize that after three months of training, infants in the music group would show more improvement than their counterparts in the control group in their (1) sensitivity to expressive features in music, and (2) general ability to relate others' action to their emotional states (i.e. emotion–action understanding), independent of music.

To measure sensitivity to expressive features in music, we used an intermodal violation-of-expectation paradigm[1] in which the infants were presented with emotionally consistent (happy and sad music with happy and sad faces, respectively) and inconsistent displays (happy and sad music with sad and happy faces, respectively) in alternate test trials. We were interested in the looking time difference between the consistent and inconsistent displays, which indicated how much the infants were surprised by the music–face inconsistency against a consistent baseline after training. We also tested emotion–action understanding with a violation-of-expectation procedure in which an agent emoted either positively or negatively (counterbalanced across infants), via facial and vocal expressions, towards one of two objects, followed by her later approaches to them (Phillips *et al.*, 2002; Vaish & Woodward, 2010). Emotion–action understanding was operationalized as the longer looking times when the agent approached the object that: (a) had not received a positive expression, compared to the other object, in the positive expression condition; (b) had received a negative expression, compared to the other object, in the negative expression condition. Consistent with previous infant research (e.g. Luo, 2011; Luo & Baillargeon, 2010; Wellman, Phillips, Dunphy-Lelii & LaLonde, 2004), we took the infants' longer looking time as a measure of their surprise response to expectancy violations. It has been suggested that to measure surprise infants' facial reactions may not be as reliable as gazing because infants show the prototypical surprise face (i.e. brow raising and mouth opening) even in non-surprising situations (Camras, 1992; Camras, Lambrecht & Michel, 1996), and they do not necessarily produce such surprise expressions (but do look longer at the target object) in unexpected events (Camras, Meng, Ujiie, Dharamsi, Miyake *et al.*, 2002; Scherer, Zentner & Stern, 2004).

[1] Our procedure differs from preferential looking in which two emotional faces are presented side by side on a screen on each trial, one being consistent while the other inconsistent with a simultaneous auditory stimulus (e.g. Flom & Whiteley, 2014; Nawrot, 2003). If infants understand the heard stimulus they are assumed to prefer and thus look longer at the face that matches it. In contrast, in the present procedure an expectation of music–face consistency was first induced via three familiarization trials which preceded the test trials and showed only music–face consistent displays. In the test trials in which only one face was shown at a time, any music–face inconsistency would violate such consistency expectation and bring about surprise indexed by increased looking times. The expectation element was further reinforced by the fact that in both the familiarization and test trials the music was turned on 10 seconds before the visual display, allowing ample time for the expectation to form. Hence, the current procedure differs from preferential looking because it is based on infant expectation, not preference (only one face is shown on each trial).

Method

Participants

We recruited typically developing 15-month-olds (plus or minus two weeks) through advertisements posted on local online parent-child forums. Eighty-two infants and their parents signed up to participate in a weekly emotion training class spanning three months at the Centre for Developmental Psychology, the Chinese University of Hong Kong. Six infants did not manage to complete the pre-test due to fussiness (1) and crying (5), and were thus excluded from the study. Two independent observers, who were blind to the experimental conditions, saw the infant's face on a monitor in a separate room and judged on the spot whether to end testing with fussy and crying infants. The remaining infants who completed the pre-test were randomly assigned to two training groups, with 40 and 36 infants in the music and control group, respectively. Nine infants dropped out during the three-month training period. Data from another three infants were discarded because they did not attend the training regularly (participating in less than 75% of the classes, n = 2) or failed to complete the post-test (1). Hence, the final sample at post-test comprised 64 18-month-olds, with 33 infants (18 boys) in the music and 31 infants (16 boys) in the control condition. Their mean ages at the pre- and post-test were 15.0 months (SD = 0.3 month) and 18.2 months (SD = 0.3 month), respectively. All the participants were ethnically Chinese and were brought up in Cantonese-speaking families from lower-middle-class neighbourhoods. The majority (84%) of them did not have siblings. There were no parental reports of perceptual, psychological or linguistic abnormalities in the infants, who had not been exposed to any regular music training prior to the pre-test. The experimental procedure was approved by the Social Science Ethics Committee, the Chinese University of Hong Kong. Parents' written informed consent on infants' participation was obtained prior to testing and training. The infants and their parents were rewarded with a certificate for their participation.

Materials and procedure

The infants and parents visited the laboratory on a weekly basis for 14 consecutive weeks, 12 visits for the training and two visits for the pre- and post-test,

Table 1 *Details of the selected musical excerpts*

Emotion	Work	Composer	Measures	Key/Mode	Pitch range (Hz)	Tempo (bpm*)
Happy	Carnaval des Animaux (Finale)	Saint-Saëns	10–26	C major	196.00– 3135.96	220
Happy	Tombeau de Couperin (Rigaudon)	Ravel	1–9	C major	220.00– 1975.53	120
Sad	Nocturne Op. 9 No. 1	Chopin	0–4	B flat minor	466.16– 1108.73	100
Sad	Kol Nidrei	Bruch	9–11	D minor	220.00– 293.66	20

*beats per minute

respectively. After completing the pre-test, the infants and their parents were randomly assigned to the music or control condition. Identical testing was administered to the two groups. The only difference lay in the training, whether or not the emotion training activities were supported by interactive music-making.

Pre-test

Parental demographic questionnaire and warm-up activities. Upon arrival on their first visit, the parents were asked to fill in a questionnaire on the basic demographics of the family and information about daily parent–child interaction (e.g. language for everyday interaction, infants' involvement in musical activities). Two experimenters played with the infant for 15 to 20 minutes to familiarize him/her with the laboratory setting. The chief experimenter then briefly introduced the experimental setup and explained the testing procedure to the parent before the tests began.

Sensitivity to expressive features in music. The infant was seated on his/her parent's lap and faced a display booth (180 cm high × 116 cm wide × 78 cm deep) with a stage (40 cm high × 108 cm wide) in the front concealed by a curtain. We mounted a digital camcorder below the stage to videotape infant behaviour during the entire procedure from an angle optimal for coding. This visual output was shown on a monitor in another room where two independent observers recorded the infants' attention to the display. The observers were adequately trained so that they were able to differentiate on the spot whether the infant was looking at the experimental display. The observers, who were naïve to the type of test events to which the infant was exposed, were asked to keep depressing a button when the infant was looking at the display and release it when the infant looked away. Time lengths of these button presses were relayed to a computer program which then worked out the infants' looking times. The looking times recorded by the primary observer were used to determine when a trial had ended, while the judgements of the secondary observer were used to calculate inter-observer reliabilities. To minimize parental interference, the parents listened to masking music via headphones and refrained from interacting with their infants throughout the procedure.

The music we used was piano renditions of the excerpts used in Peretz, Gagnon and Bouchard (1998) and Terwogt and van Grinsven (1991). These excerpts are from the Western classical repertoire and could communicate happy and sad feelings. The piano renditions were played and recorded by a graduate student in music. Thirty-two adults rated these piano excerpts on how happy and sad they were, using a 9-point Likert scale. The two happy and two sad excerpts that received the most extreme happy and sad ratings, respectively, were selected and used in the familiarization and test trials. The happy excerpts were in the major mode and played at a fast tempo between 120 and 220 beats per minute (bpm), lying in the medium to high pitch range. The sad excerpts were in the minor mode and played in a slow tempo varying from 20 to 100 bpm. The pitch range was relatively restricted. The pedal was used in making the sad excerpts. Details of the structural features of these piano renditions are shown in Table 1. They were digitally recorded and presented via loudspeakers at 45–55 dB (infant ear level) against a background of 20 dB.

Two amateur actresses were recruited to pose a happy and a sad expression. In both emotions, the actress was videotaped from her shoulders up against a white wall while simulating the following episode in Cantonese in a happy or sad manner: 'Here's your milk. Drink it all up. Don't spill any. Oh look, you got some on the tray. Don't play in it. I'll get a sponge and clean it up.' We used this script because young children could discriminate between a happy and sad rendition of it even when muted (Stifer & Fox, 1987). The actresses were trained to demonstrate the facial features that defined the respective emotions in accordance with the Facial Action Coding System (Ekman & Friesen, 1975). In the happy expression, the actress smiled or grinned throughout with the corners of her lips lifted and the skin below her lower eyelid pushed up. She also showed naso-labial folds and her cheek was consistently raised. In the sad expression, the corners of the actress's lips were down and her lips occasionally trembled. The inner corners of her eyebrows and upper eyelids were drawn up resulting in a triangulated shape in each of her brows. Sometimes her eyes cast down slightly.

Testing consisted of three familiarization and four subsequent test trials. Familiarization started when a musical excerpt (happy or sad; counterbalanced across infants) was played to the infant. After infants had listened to the music for 10 seconds, the closed curtain was drawn revealing a 24-inch monitor on which a silent video clip was played, showing an actress posing either a happy or sad expression. The monitor was 90 cm away from the infant. The music kept playing while the video was shown, and the emotion in the music always matched the emotion expressed by the actress in the video. The trial ended when the infant looked away for 2 consecutive seconds after having looked for at least 5 cumulative seconds, or when 30 cumulative seconds had elapsed. The ensuing test trials followed the same procedure, except that another actress was shown in the video and another set of happy and sad music was played. The emotion in the music was either consistent or inconsistent with the emotion expressed by the actress, and there were two consistent and two inconsistent test trials. A test trial ended when the infant looked away for 2 consecutive seconds or looked for 60 cumulative seconds without looking away. The consistent and inconsistent trials alternated and their presentation order was counterbalanced across infants. Presentation order of the happy and sad expressions was also counterbalanced. The test took about 10 minutes.

Emotion–action understanding. The infants sat on their parents' lap facing the stage, on which there placed two identical toy bears in different colours laid 48 cm apart. Infant behaviour was filmed by a camcorder placed below the stage between the two toy bears. Infant looking times were recorded in the same way as in the previous test by two trained observers. Another camcorder was mounted behind and above the infant to videotape the experimental events. The audio-visual output was shown on a computer screen online in the control room where the chief experimenter could decide on the spot whether the actress was following the scripts and displaying the appropriate facial and vocal expressions, and whether she was acting consistently in the test events across the experimental conditions. Trials on which these requirements were not met were marked to be discarded later. This audio-visual output was also recorded for further offline rating by independent raters for acting consistency in the test events across the experimental conditions (see Reliability coding). Throughout the testing procedure, parents were instructed not to interact with their infants, either verbally or via body movement, so as to minimize parental interference.

The infants were randomly assigned either to a happy or disgust condition, each comprising three familiarization and four test trials. Familiarization began when the curtain was drawn showing an actress sitting behind a stage on which two toy bears in different colours were resting in front of the actress. In the happy condition, after the infant looked at the display for 2 seconds, the actress smiled and looked at one of the two bears, saying 'Oh! Look at that teddy bear!' in a pleasant voice filled with excitement and interest. After 5 seconds the curtain was closed independent of the infant's response. When the curtain reopened, the actress held and looked down at the same bear that she had previously emoted positively about. The trial ended when the infant looked away from the display for 2 consecutive seconds, or 30 cumulative seconds had elapsed.

Four test events (two consistent and two inconsistent) were administered to each infant after familiarization. In a consistent test event, the pre-trial began with the actress looking at and emoting positively towards the other bear, the one not regarded or held during familiarization. The curtain was then closed. After 1 second the curtain reopened and the main trial began with the actress smiling at and grasping this same bear. This event sequence was consistent because in the main trial the actress approached the bear that she had emoted with positive facial and vocal expressions in the pre-trial, thus demonstrating consistency between her expressed emotion and subsequent action towards this object. In an inconsistent event, in contrast, the actress emoted positively in the pre-trial towards the bear she had held during familiarization, but subsequently held the other bear in the main trial. This event sequence was inconsistent because the actress did not grasp the object that she previously felt excited about.

In the disgust condition, in familiarization, the actress lowered and drew her brows together, retracted her neck and said 'Yuck! Look at that teddy bear!' in a disgusted tone of voice towards one of the two bears (Bear A). After the curtain reopened, the actress regarded and held the other bear which was not attended to seconds before (Bear B). Four test events (two consistent and two inconsistent) followed familiarization. In a consistent test event, the pre-trial began when the actress emoted negatively towards the bear that had not received a negative expression during familiarization (i.e. Bear B), and subsequently approached the other bear (i.e. Bear A) in the main trial after the curtain reopened. In contrast, the actress in an inconsistent event emoted negatively in the pre-trial towards the bear she had regarded negatively in familiarization (i.e. Bear A), but later grasped this same bear in the main trial. The main trial ended when the infant looked away from the display for 2 consecutive seconds, or 30 (familiarization trials) or 60 (test trials) cumulative seconds had elapsed. Presentation order of the consistent and inconsistent trials, the locations of the two toy bears, and the bear grasped in familiarization were counterbalanced across infants. This test took about 10 minutes and a short break was given in between the two tests.

Early communicative development. The parents were asked to complete the Cantonese version of the MacArthur-Bates Communicative Development Inventory (CCDI; Fenson, Dale, Reznick, Bates, Thal *et al.*, 1994). This inventory is a standardized parental report for assessing infants' and toddlers' language and communicative skills with excellent internal reliability (Cronbach's alpha: Level 1: 0.97; Level 2: 0.99; Fenson *et al.*, 1994). Because the infants at the pre-test were 15 months of age, the infant form was used (Level 1: 8 to 18 months), which primarily measured the use of gestures and actions, and vocabulary comprehension and production. We administered the short form because it was more time-efficient and was adequately correlated with the long form ($r = .98$ for vocabulary comprehension, and $r = .97$ for vocabulary production; Fenson, Pethick, Renda, Cox, Dale *et al.*, 2000). The maximum score for the inventory is 205.

The training protocol

Music condition. Training in the music condition followed the Suzuki principles of parental involvement, learning from observing peers, and repetition in the Early Childhood Education programme (Jones, 2004). The infants and parents assigned to this condition were divided into groups of five to six infant-parent dyads and participated in a series of interactive musical activities designated to illustrate different basic emotions. In the weekly 45-minute classes, a trained experimenter (i.e. the teacher), who was a university-trained professional experienced in teaching infants and toddlers, led the dyads to sing, move around, and play various toy percussion instruments to build a repertoire of nursery rhymes, action songs, and lullabies. Specifically, in the first six weeks, training started with songs and dances that introduced some basic facial features and body parts for expressing emotions. In the remaining weeks, the teacher demonstrated the emotions of happiness, sadness, fear, anger, disgust, and surprise by showing the respective facial and body movements in response to songs conveying these emotions. The training also included story-telling sessions in which the teacher posed emotionally congruent facial and bodily expressions when she was telling the story, coupled with background music and singing. All these activities were supported by music-making (e.g. singing, dancing, or playing toy percussion instruments) or passive exposure to background music. The parents participated actively to engage their infants in the activities. The infant-parent dyads were also encouraged to interact with one another and to repeat such interaction during class so that the infants could learn from observing and imitating their peers.

Control condition. We included a control condition to examine the possibility that it was the highly interactive nature of the training itself, rather than the engagement in music, that fostered the infants' affective interpretation of behaviour. The infants and parents assigned to the control condition received similar interactive training on the six basic emotions, but without any form of musical engagement. There was no singing, dancing, percussion instrument playing, nor passive listening to music. To introduce the facial and body parts that communicate emotions, the teacher used a puppet to prompt the infant-parent dyads to touch the corresponding areas. The teacher also displayed different facial and bodily expressions to teach the basic emotions in the story-telling sessions. We also arranged other equally interactive activities such as playing with blocks and balls and making handicrafts to engage the infants. None of the activities were accompanied by background music and music-making.

Post-test

Infants from both groups underwent the same test procedure at the post-test, which involved the same two paradigms as in the pre-test. The toddler form (Level 2: 16 to 30 months) of the MacArthur-Bates Communicative Development Inventory was used to measure communicative development because it better suited their age (18 months). The toddler form is more focused on vocabulary production and combination of words; the maximum score is 138. The parents also reported the numbers of hours that their infants were engaged in music and talking at home during the training period. Infants from both groups had similar amounts of musical (music group: $M = 2.96$ hours per day, $SE = .54$; control group: $M = 2.05$ hours per day, $SE = .35$; $p = .128$) and verbal interaction (music group: $M = 2.66$ hours per day, $SE = .45$; control group:

M = 2.61 hours per day, SE = .54; p = .941) with their parents. Hence the parents from the music group were not more likely than those from the control group to engage their infants musically at home in the training period.

Reliability coding

To assess inter-observer agreement in recording the infants' looking times, each familiarization and test trial was divided into 100-ms intervals and a computer program tracking infant looking times compared the two observers' binary judgements on whether the infant was looking at the display for each of these intervals. Overall inter-observer reliability was calculated by dividing the number of intervals in which the two observers agreed by the total number of intervals in a trial. Recordings with reliabilities lower than .80 were to be discarded due to judgement inconsistency between the two observers. The overall reliability at the pre- and post-test was .91 and .94, respectively. We also examined the extent to which the two observers agreed to end a trial by computing two additional indices. First, a separate inter-observer reliability was calculated using the above method for the last 2 seconds of each trial; this reliability averaged .96 for both the pre- and post-test. Second, a trial-end reliability was calculated for each trial by dividing the number of agreed units after the last disagreed unit within the last 2 seconds by 20 (i.e. total number of units during the last 2 seconds); this reliability averaged .92 and .95 for the pre- and post-test, respectively. Hence, the two observers were in excellent agreement regarding their decisions to end a trial.

The actress's acting in the emotion–action understanding test events was evaluated offline to confirm that it did not vary systematically across the experimental conditions. Two naïve raters watched videos showing the actress's acting in both the pre- (emoting before grasping) and main trials (grasping), and were asked to decide whether the acting was from the consistent or inconsistent events. For the pre-trials, overall accuracies were 53% and 47% for the consistent and inconsistent events, respectively; for the main trials, overall accuracies were 52% and 55% for the consistent and inconsistent events, respectively. These results indicated near-chance discrimination. The raters were also asked to rate on the intensity of happiness or disgust displayed in each video using a 9-point scale. The perceived emotional intensities were comparable across the consistent (average rating = 8.1) and inconsistent events (average rating = 7.8). Hence, actress bias in enacting the events appeared to be minimal.

Results

No data were discarded because of inappropriate acting and overall inter-observer reliabilities lower than .80. Preliminary examination revealed no significant differences in any of the demographic variables and the quality of parent–child interaction (i.e. amount of verbal and musical interaction at home in the training period) between the music and control group. The two groups did not differ in the MacArthur-Bates CDI scores before, $t(62) = -.30$, p = .766, and after the training, $t(62) = .15$, p = .885.

Sensitivity to expressive features in music

Preliminary analyses showed no significant correlations between the infants' looking times and gender, socioeconomic status, the emotion first shown in familiarization, and the order of test trials. Hence these variables were collapsed in all subsequent analyses.

At the pre-test, the mean looking times for the first, second, and third familiarization trials were 26.36s (SD = 7.36s), 24.18s (SD = 8.15s), and 22.01s (SD = 10.35s), respectively. These looking times were submitted to a repeated measures analysis of variance (ANOVA), showing a significant trend of decrement from the first to the third trial, $F(2, 62) = 10.36$, $p < .001$. At the post-test, the mean looking times in familiarization similarly declined from 26.94s (SD = 6.45s) to 22.39s (SD = 9.59s). An ANOVA confirmed that the decrement was significant, $F(2, 62) = 6.37$, p = .003. These results suggested that the infants at the pre- and post-tests were actively attending to the information presented in familiarization and were gradually habituated to it.

Infants' mean looking times in the test trials at the pre- and post-tests are presented in Figure 1. These looking times were subjected to a 2 × 2 × 2 mixed-design ANOVA with time (pre-test vs. post-test) and trial type (consistent vs. inconsistent trials) being the within-subject factors and group (music vs. control) being the between-subjects factor. The only significant effect was the three-way interaction, $F(1, 62) = 10.08$, p = .002. To understand this effect, we conducted a Time × Trial Type ANOVA separately for the music and control groups. The only significant effect was the two-way interaction in the music group, $F(1, 32) = 10.35$, p = .003 (control group: $F(1, 30) = 1.74$, p = .197). At the pre-test, the infants in neither group showed significant looking time differences between the consistent and inconsistent trials: music, $t(32) = 1.51$, p = .142; control, $t(30) = .30$, p = .769. At the post-test, however, the infants in the music group looked significantly longer in the inconsistent (M = 23.32s, SD = 15.29s) than consistent trials

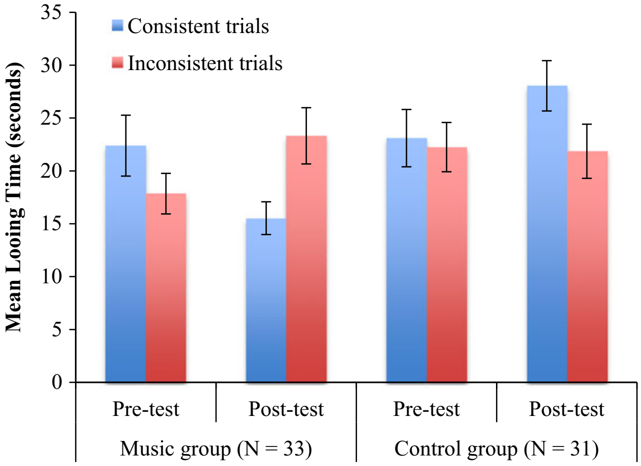

Figure 1 *Infants' average looking times in the sensitivity to expressive features in music test.*

(M = 15.53s, SD = 8.90s), $t(32) = 2.90$, p = .007, whereas those in the control group looked longer in the consistent (M = 28.05s, SD = 13.23s) than inconsistent trials (M = 21.86s, SD = 14.27s), $t(30) = -2.10$, p = .052.

We also examined within each group the change in looking times across the two time points. For the music group, looking times in the consistent trials dropped from the pre- (M = 22.39s, SD = 16.54s) to post-test (M = 15.53s, SD = 8.89s), $t(32) = 2.14$, p = .040, whereas those in the inconsistent trials increased from 17.85s (SD = 11.02s) to 23.32s (SD = 11.04s) across the time points, $t(32) = -1.94$, p = .058. In contrast, the control infants' looking times in both the consistent and inconsistent trials did not change significantly after the training: consistent, $t(30) = -1.38$, p = .178; inconsistent, $t(30) = .13$, p = .900.

We computed a difference score for each infant at each time point by taking the looking time difference between the consistent and inconsistent trials. We assumed that this score indicated the infants' awareness of how much the actress's facial expression matched the music, thus suggesting sensitivity to expressive features in music. An ANOVA revealed a Time × Group interaction on such sensitivity, $F(1, 62) = 10.08$, p = .002, d = 0.40. For the music group, the looking time difference at the post-test was significantly greater than that at the pre-test, $t(32) = -3.22$, p = .003. For the control group this looking time difference did not change significantly across the time points, $t(30) = 1.32$, p = .197. The results suggested that the music group improved more than the control group on sensitivity to musical expression after the emotion training.

Emotion–action understanding

Preliminary analyses revealed no relationships between looking times and gender, socioeconomic status, emotion condition (i.e. happy vs. disgusted), locations of target objects, the target grasped in familiarization, and the order of test trials. These dimensions were therefore collapsed in all subsequent analyses.

A repeated measures ANOVA revealed a significant overall decrease in looking time across the three familiarization trials at the pre-test from 24.84s (SD = 7.56s) to 18.13s (SD = 9.09s), $F(2, 62) = 27.71$, $p < .001$. A similar decrement was also evident at the post-test, from 26.09s (SD = 7.34s) to 18.57s (SD = 9.66s), $F(2, 62) = 14.44$, $p < .001$, suggesting that the infants encoded the event sequences and were gradually habituated to them in familiarization.

Infants' mean looking times in the test trials at both time points are presented in Figure 2. These data were analysed by a 2 × 2 × 2 (Times × Trial Types × Groups) mixed-design ANOVA. The significant effects were the three-way interaction, $F(1, 62) = 4.93$, p = .030, Trial Type × Group interaction, $F(1, 62) = 4.76$, p = .033, and Time × Trial Type interaction, $F(1, 62) = 4.32$, p = .042. To interpret the three-way interaction, we examined the Time × Trial Type effect separately for the music and control groups. The Time × Trial Type interaction was significant in the music group, $F(1, 32) = 7.97$, p = .008, but not in the control group, $F(1, 30) = .01$, p = .912. At the pre-test, the infants in both groups had similar looking times in the consistent and inconsistent trials: music, $t(32) = .50$, p = .619; control, $t(30) = 1.13$, p = .269. Nevertheless, at the post-test the music group looked significantly longer in the inconsistent (M = 18.51s, SD = 11.84s) than consistent trials (M = 12.98s, SD = 8.98s), $t(32) = 2.59$, p = .014. This effect was not present for the control group: consistent,

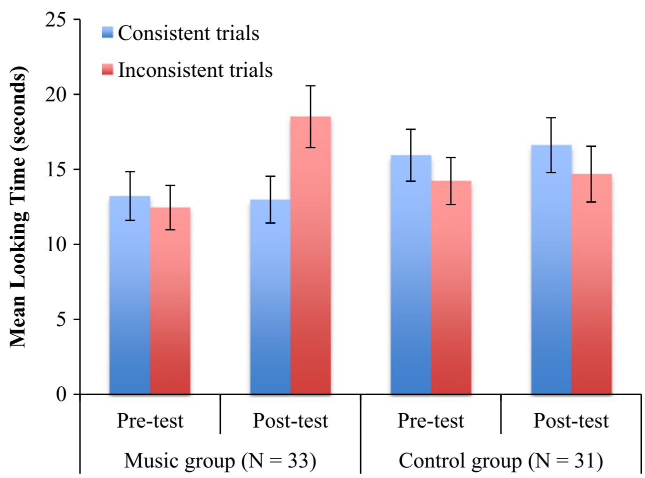

Figure 2 *Infants' average looking times in the emotion–action understanding test.*

M = 16.61s, SD = 10.16s; inconsistent, M = 14.68s, SD = 10.34s; $t(30)$ = 1.28, p = .210. For the music group, looking times in the inconsistent trials increased from the pre- (M = 12.45s, SD = 8.48s) to post-test (M = 18.51s, SD = 11.83s), $t(32) = -2.61$, p = .014, but those in the consistent trials remained unchanged across the time points, $t(32)$ = .11, p = .913. The control infants' looking times in both types of test trials did not change after the training: consistent, $t(30) = -.20$, p = .843; inconsistent, $t(30) = -.30$, p = .767.

We also compared the increments in looking time differences (inconsistent – consistent) from the pre- to post-test for the two groups to evaluate the training effect. ANOVA results showed a significant Time × Group interaction, $F(1, 62)$ = 4.93, p = .030, d = 0.29. The increment was significant for the music, $t(32) = -2.83$, p = .008, but not the control group, $t(30)$ = .11, p = .912. Hence, the infants who had participated in emotion training supported by musical activities were more able to use the actress's facial and vocal expressions to predict her subsequent grasping action, compared to those who had received the same emotion training without musical engagement.

Discussion

The present findings demonstrate that experiencing emotion in music-making fosters (a) 18-month-olds' sensitivity to expressive features in music and (b) their affective interpretation of behaviour. Recent evidence has indicated that musically trained infants, but not their untrained counterparts, preferred listening to music that conformed to Western tonal and metrical structures (Gerry, Faux & Trainor, 2010; Gerry *et al.*, 2012), suggesting an impact of musical experience on perceiving pitch- and rhythm-related cues early in infancy. Using a different paradigm, we found that infants who received emotion training coupled with musical activities showed an increased sensitivity to expressions in music as indicated by greater novelty responses to music–face inconsistent displays. The infants from the music group tended to look longer at the emotionally discordant displays and less at the concordant ones at the post- than pre-test, yet this was not observable in the control group. We thus think that the repeated pairing of music and affectively consistent facial–vocal displays in the training exerted two different effects on the music group infants over time. First, after such experience of music–face consistency, any music–face match at the post-test would appear to be well expected or unexceptional. This contributed to the shortened looking times in the consistent trials via a process similar to habituation. Second, as hypothesized, training also rendered any music–face mismatch at the post-test particularly surprising and thus contributed to the increased looking times in the inconsistent trials. This also explains why the control infants' attention to the displays did not change significantly across the two time points, possibly because they had not experienced such music–face consistency in their training. These results converge with the recent evidence pointing to an early effect of music training on infants' sensitivity to auditory features in music (Gerry *et al.*, 2010, 2012).

The literature has established that young children are responsive to contrasting acoustic patterns in their interpretation of the emotional content of music. For instance, 4- to 5-year-olds use tempo to categorize music as happy or sad (Dalla Bella *et al.*, 2001; Mote, 2011); children at age 6 to 8 start using the major–minor mode distinction to judge the emotion behind music (Dalla Bella *et al.*, 2001; Gerardi & Gerken, 1995; Gregory, Worrall & Sarge, 1996). Our findings indicate that infants at 18 months with regular music engagement are able to detect the inconsistency between music and facial displays in emotional terms. Therefore, it is possible to promote infants' emerging sensitivity to features conveying affective meaning in music via regular music-making. We consider such responsiveness to inconsistent music–face expressions to be a precursor to developing a full-fledged understanding of the emotional implications of music.

As discussed, our analyses of the music test data showed that for the music group, looking times decreased in the consistent but increased in the inconsistent trials from the pre- to post-test. Yet these changes over time were not observable in the control group. Looking from this perspective, training appeared to have exerted no noticeable effects on the control infants' sensitivity to expressive features in music. However, for these infants the looking time difference between the consistent and inconsistent trials at the post-test reached marginal significance unexpectedly (p = .052), with longer looking times from the consistent trials. We think that this marginally significant effect may be understood with reference to a broader observation, which is as follows. In fact, infants in both groups tended to look longer at the consistent than inconsistent events at the pre-test, although these differences did not reach significance. What is more, a similar insignificant trend is also observable in the emotion–action understanding test for both the music (pre-test) and control infants (pre- and post-test). This raises the interesting possibility that at 15 and 18 months of age infants may already have an emerging sense of how facial/vocal expressions are naturally associated with music expressions and action, so that they are initially more inclined and attracted toward events in which this association is preserved (i.e. the consistent events). But note that this trend may not be reliable enough to support statistical significance in any consistent way with the current sample size and paradigm. On this basis, as discussed, the repeated pairing of music and affectively consistent facial–vocal displays in the current training habituated the music group infants to this association and therefore looking times decreased in the consistent trials. Training also rendered any music–face mismatch at the post-test particularly surprising and thus looking times increased in the inconsistent trials. This may explain why only the music group infants at the post-test looked longer at the inconsistent than consistent events, whereas in all the other conditions in both tasks the opposite seemed to be observable.

More importantly, results from the emotion–action understanding test showed that music-making in a social context promoted 18-month-olds' understanding of the general relationship between emotion and action. The experience and learning of emotion in musical activities hone infants' ability to use others' emotional expressions to interpret their actions. Our findings are consistent with the recent claim that engagement in artwork enhances cognitive and affective theory-of-mind (ToM) (Goldstein & Winner, 2012; Kidd & Castano, 2013). The prevailing explanation for this effect has to do with the need to deploy ToM in creating and appreciating art. We engage ourselves in the arts to release or express our inner feelings and ideas. Exposure to art prompts us to become more aware of our own and others' thoughts and emotions because this is how we make sense of the art creation and appreciation process. This inevitably draws on the same pool of socio-cognitive resources that also supports ToM. Our study extends the existing evidence to music via an experimental approach and demonstrates that interaction with music leads to heightened sensitivities to musical expressions and the general emotion–action relationship. Music is a powerful medium that is capable of eliciting strong emotions in the audience (Juslin & Laukka, 2003; Sloboda & Juslin, 2001). Hence, we think that its effect may be more specific to promoting the understanding of emotion than other mental states. It would be particularly interesting to examine in future studies the impact of music-making on the representation of non-affective mental states such as intention and belief, and to compare this with music's effect on the affective aspect of mentalistic reasoning.

While both the music and control training programmes in the present study are assumed to be equally interactive, it is plausible that the synchronized movements that are unique to the music programme have contributed the most to enhancing emotion–action understanding. It has been demonstrated that motor synchrony directs our attention towards the interacting partners (Macrae, Duffy, Miles & Lawrence, 2008) and leads us to consider one another as similar (Valdesolo, Ouyang & DeSteno, 2010; Valdesolo & DeSteno, 2011). Motor synchrony also elicits more prosocial behaviours in infants (Cirreli *et al.*, 2014) and adults (Valdesolo & DeSteno, 2011). We argue that the interpersonal synchrony in music-making, via a cooperation-enhancing mechanism, may also motivate a representation of others' emotion when we try to affiliate with them. In fact this conjecture is also relevant to interpreting the positive impact of group music-making on 12-month-olds' social-emotional development (Gerry *et al.*, 2012) and 4-year-olds' prosociality (Kirschner & Tomasello, 2010). Infants and children may perceive a sense of unity and connectedness with their partners in coordinated activities in music, thus engaging themselves in more social cohesive behaviours. Further investigation is needed to delineate the exact relationships among synchronized action in music-making, representation of emotional states, and actual prosocial behaviours.

Overall, our study provides an empirical demonstration that interactive music-making, a ubiquitous form of arts-making, enhances 18-month-olds' sensitivity to musical expressions and their emotion–action understanding. Given that the infants from the control group were exposed to similar engaging activities but without music, we are able to argue that music has a unique effect on infants' affective interpretation of behaviour in addition to teaching and learning based on verbal communication. To prelinguistic infants and their caregivers, music provides an effective avenue by which emotions can be expressed and communicated. Our findings suggest that simple music activities could be introduced early in development to promote infants' emotional and socio-cognitive development. Specifically, an experience with music benefits infants' representation of emotional cues that help them understand others' behaviours, an ability fundamental to successful social interaction.

Acknowledgements

This research was funded by a Direct Grant for Research to the corresponding author from the Social Science Panel, the Chinese University of Hong Kong (project number 4052003). We are grateful to Maggie Ho and Wendy Liu for hosting the emotion understanding classes. We thank all the infants and parents for participating in the research. We also thank Eric Chan, Esther Chan, Gloria Chan, Lawrence Chan, Rachel Chuang, and Sally Tong for assistance in testing the infants, and two anonymous reviewers for their valuable suggestions in improving the manuscript.

References

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.

Balkwill, L., Thompson, W.F., & Matsunaga, R. (2004). Recognition of emotion in Japanese, Western, and Hindustani music by Japanese listeners. *Japanese Psychological Research*, **46**, 337–349.

Barna, J., & Legerstee, M. (2005). Nine- and twelve-month-old infants relate emotions to people's actions. *Cognition and Emotion*, **19**, 53–67.

Bhatara, A., Tirovolas, A.K., Duan, L.M., Levy, B., & Levitin, D.J. (2011). Perception of emotional expression in musical performance. *Journal of Experimental Psychology: Human Perception and Performance*, **37**, 921–934.

Bigand, E., & Poulin-Charronnat, B. (2006). Are we 'experienced listeners'? A review of the musical capacities that do not depend on formal musical training. *Cognition*, **100**, 100–130.

Brune, C.W., & Woodward, A.L. (2007). Social cognition and social responsiveness in 10-month-old infants. *Journal of Cognition and Development*, **8**, 133–158.

Budd, M. (1985). *Music and the emotions: The philosophical theories*. London: Routledge.

Camras, L.A. (1992). Expressive development and basic emotions. *Cognition and Emotion*, **6**, 269–284.

Camras, L.A., Lambrecht, L., & Michel, G. (1996). Infant 'surprise' expressions as coordinative motor structures. *Journal of Nonverbal Behaviour*, **20**, 183–195.

Camras, L.A., Meng, Z., Ujiie, T., Dharamsi, S., Miyake, K., et al. (2002). Observing emotion in infants: facial expression, body behavior, and rater judgments of responses to an expectancy-violating event. *Emotion*, **2**, 179–193.

Cannon, E.N., & Woodward, A.L. (2012). Infants generate goal-based action predictions. *Developmental Science*, **15**, 292–298.

Cannon, E.N., Woodward, A.L., Gredeback, G., von Hofsten, C., & Turek, C. (2012). Action production influences 12-month-old infants' attention to others' actions. *Developmental Science*, **15**, 35–42.

Cicchino, J.B., & Rakison, D.H. (2008). Producing and processing self-propelled motion in infancy. *Developmental Psychology*, **44**, 1232–1241.

Cirelli, L.K., Einarson, K.M., & Trainor, L.J. (2014). Interpersonal synchrony increases prosocial behavior in infants. *Developmental Science*, **17**, 1003–1011.

Conrad, N.J., Walsh, J., Allen, J.M., & Tsang, C.D. (2011). Examining infants' preferences for tempo in lullabies and playsongs. *Canadian Journal of Experimental Psychology*, **65**, 168–172.

Cooke, D. (1959). *The language of emotion*. Oxford: Oxford University of Press.

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.

Davies, S. (2001). Philosophical perspectives on music's expressiveness. In P.N. Juslin & J.A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 23–44). New York: Oxford University Press.

Dissanayake, E. (2000). Antecedents of the temporal arts in early mother–infant interaction. In N.L. Wallin, B. Merker & S. Brown (Eds.), *The origins of music* (pp. 389–410). Cambridge, MA: The MIT Press.

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.

Ekman, P., & Friesen, W. (1975). *Unmasking the face*. Englewood Cliffs, NJ: Prentice Hall.

Falk, D. (2004). Prelinguistic evolution in early hominins: whence motherese? *Behavioral and Brain Sciences*, **27**, 491–503.

Fenson, L., Dale, P.S., Reznick, J.S., Bates, E., & Thal, D. et al. (1994). Variability in early communicative development.. *Monographs of the Society for Research in Child Development*, **59** (5, Serial No. 242).

Fenson, L., Pethick, S., Renda, C., Cox, J., Dale, P.S., et al. (2000). Short-form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics*, **21**, 95–116.

Fernald, A. (1991). Prosody in speech to children: prelinguistic and linguistic functions. *Annuals of Child Development*, **8**, 43–80.

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

Flom, R., & Pick, A.D. (2012). Dynamics of infant habituation: infants' discrimination of musical excerpts. *Infant Behavior and Development*, **35**, 697–704.

Flom, R., & Whiteley, M.O. (2014). The dynamics of intermodal matching: seven- and 12-month-olds' intermodal matching of affect. *European Journal of Developmental Psychology*, **11**, 111–119.

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

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

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

Gerry, D.W., Faux, A.L., & Trainor, L.J. (2010). Effects of Kindermusik training on infants' rhythmic enculturation. *Developmental Science*, **13**, 545–515.

Gerry, D., Unrau, A., & Trainor, L.J. (2012). Active music classes in infancy enhance musical, communicative and social development. *Developmental Science*, **15**, 398–407.

Goldstein, T.R., & Winner, E. (2012). Enhancing empathy and theory of mind. *Journal of Cognition and Development*, **13**, 19–37.

Gredeback, G., & Melinder, A. (2010). Infants' understanding of everyday social interactions: a dual process account. *Cognition*, **114**, 197–206.

Gregory, A.H., Worrall, I., & Sarge, A. (1996). The development of emotional responses to music in young children. *Motivation and Emotion*, **20**, 341–349.

Hepach, R., & Westermann, G. (2013). Infants' sensitivity to the congruence of others' emotions and actions. *Journal of Experimental Child Psychology*, **115**, 16–29.

Husain, G., Thompson, W.F., & Schellenberg, E.G. (2002). Effects of musical tempo and mode on arousal, mood, and spatial abilities. *Music Perception*, **20**, 151–171.

Johnson, S.P., Davidow, J., Hall-Haro, C., & Frank, M.C. (2008). Development of perceptual completion originates in information acquisition. *Developmental Psychology*, **44**, 1214–1224.

Jones, D. (2004). Suzuki early childhood education. *American Suzuki Journal*, **36**, 32–38.

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

Juslin, P.N., & Laukka, P. (2004). Expression, perception, and induction of musical emotions: a review and a questionnaire study of everyday listening. *Journal of New Music Research*, **33**, 217–238.

Kidd, D.C., & Castano, E. (2013). Reading literary fiction improves theory of mind. *Science*, **342**, 377–380.

Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior*, **31**, 354–364.

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

Libertus, K., & Needham, A. (2011). Reaching experience increases face preference in 3-month-old infants. *Developmental Science*, **14**, 1355–1364.

Loucks, J., & Sommerville, J.A. (2012). The role of motor experience in understanding action function: the case of the precision grasp. *Child Development*, **83**, 801–809.

Luo, Y. (2011). Three-month-old infants attribute goals to a non-human agent. *Developmental Science*, **14**, 453–460.

Luo, Y., & Baillargeon, R. (2010). Toward a mentalistic account of early psychological reasoning. *Current Directions in Psychological Science*, **19**, 301–307.

Macrae, C.N., Duffy, O.K., Miles, L.K., & Lawrence, J. (2008). A case of hand waving: action synchrony and person perception. *Cognition*, **109**, 152–156.

Meltzoff, A.N. (2007). 'Like me': a foundation for social cognition. *Developmental Science*, **10**, 126–134.

Meltzoff, A.N., & Brooks, R. (2008). Self-experience as a mechanism for learning about others: a training study in social cognition. *Developmental Psychology*, **44**, 1257–1265.

Mohn, C., Argstatter, H., & Wilker, F.W. (2010). Perception of six basic emotions in music. *Psychology of Music*, **39**, 503–517.

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

Myowa-Yamakoshi, M., Kawakita, Y., Okanda, M., & Takeshita, H. (2011). Visual experience influences 12-month-old infants' perception of goal-directed actions of others. *Developmental Psychology*, **47**, 1042–1049.

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

Needham, A. (2000). Improvements in object exploration skills may facilitate the development of object segregation in early infancy. *Journal of Cognition and Development*, **1**, 131–156.

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

Perone, S., Madole, K.L., Ross-Sheehy, S., Carey, M., & Oakes, L.M. (2008). The relation between infants' activity with objects and attention to object appearance. *Developmental Psychology*, **44**, 1242–1248.

Phillips, A.T., Wellman, H.M., & Spelke, E.S. (2002). Infants' ability to connect gaze and emotional expression to intentional action. *Cognition*, **85**, 53–78.

Rakison, D.H., & Krogh, L. (2012). Does causal action facilitate causal perception in infants younger than 6 months of age? *Developmental Science*, **15**, 45–53.

Repacholi, B.M., & Gopnik, A. (1997). Early reasoning about desires: evidence from 14- and 18-month-olds. *Developmental Psychology*, **33**, 12–21.

Rock, A.M.L., Trainor, L.J., & Addison, T.L. (1999). Distinctive messages in infant-directed lullabies and play songs. *Developmental Psychology*, **35**, 527–534.

Scherer, K.R., Zentner, M.R., & Stern, D. (2004). Beyond surprise: the puzzle of infants' expressive reactions to expectancy violation. *Emotion*, **4**, 389–402.

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

Sloboda, J.A., & Juslin, P.N. (2001). *Music and emotion: Theory and research*. Oxford: Oxford University Press.

Sommerville, J.A., Hildebrand, E.A., & Crane, C.C. (2008). Experience matters: the impact of doing versus watching on infants' subsequent perception of tool-use events. *Developmental Psychology*, **44**, 1249–1256.

Sommerville, J.A., & Woodward, A.L. (2005). Pulling out the intentional structure of human action: the relation between action production and processing in infancy. *Cognition*, **95**, 1–30.

Sommerville, J.A., Woodward, A.L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition*, **96**, 1–11.

Stifer, C.A., & Fox, N.A. (1987). Preschool children's ability to identify and label emotions. *Journal of Nonverbal Behavior*, **11**, 43–54.

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

Trainor, L.J., & Corrigall, K.A. (2010). Music acquisition and effects of musical experience. *Music Perception: Springer Handbook of Auditory Research*, **36**, 89–127.

Trehub, S.E. (2001). Musical predispositions in infancy. *Annals of the New York Academy of Sciences*, **930**, 1–16.

Trehub, S.E. (2007). Infants as musical connoisseurs. In G.E. McPherson (Ed.), *The child as musician: A handbook of musical development* (pp. 33–50). Oxford: Oxford University Press.

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

Tsang, C.D., & Conrad, N.J. (2010). Does the message matter? The effect of song type on infants' pitch preferences for lullabies and playsongs. *Infant Behavior and Development*, **33**, 96–100.

Vaish, A., & Woodward, A. (2010). Infants use attention but not emotions to predict others' actions. *Infant Behavior and Development*, **33**, 79–87.

Valdesolo, P., Ouyang, J., & DeSteno, D. (2010). The rhythm of joint action: synchrony promotes cooperative ability. *Journal of Experimental Social Psychology*, **46**, 693–695.

Valdesolo, P., & DeSteno, D. (2011). Synchrony and the social tuning of compassion. *Emotion*, **11**, 262–266.

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

Volkova, A., Trehub, S.E., & Schellenberg, E.G. (2006). Infants' memory for musical performances. *Developmental Science*, **9**, 583–589.

Wellman, H.M., Phillips, A.T., Dunphy-Lelii, S., & LaLonde, N. (2004). Infant social attention predicts preschool social cognition. *Developmental Science*, **7**, 283–288.

Woodward, A.L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, **69**, 1–34.

Woodward, A.L., & Guajardo, J.J. (2002). Infants' understanding of the point gesture as an object-directed action. *Cognitive Development*, **17**, 1061–1084.

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

Received: 4 August 2014
Accepted: 7 July 2015