Journal of Music Therapy Advance Access published December 31, 2016

Journal of Music Therapy, 00(00), 2016, 1–25
doi:10.1093/jmt/thw017

The Influence of Music on Facial Emotion Recognition in Children with Autism Spectrum Disorder and Neurotypical Children

Laura S. Brown, PhD, MT-BC

Ohio University

Background: *Children with autism spectrum disorder (ASD) often struggle with social skills, including the ability to perceive emotions based on facial expressions. Research evidence suggests that many individuals with ASD can perceive emotion in music. Examining whether music can be used to enhance recognition of facial emotion by children with ASD would inform development of music therapy interventions.*

Objective: *The purpose of this study was to investigate the influence of music with a strong emotional valance (happy; sad) on children with ASD's ability to label emotions depicted in facial photographs, and their response time.*

Methods: *Thirty neurotypical children and 20 children with high-functioning ASD rated expressions of happy, neutral, and sad in 30 photographs under two music listening conditions (sad music; happy music). During each music listening condition, participants rated the 30 images using a 7-point scale that ranged from very sad to very happy. Response time data were also collected across both conditions.*

Results: *A significant two-way interaction revealed that participants' ratings of happy and neutral faces were unaffected by music conditions, but sad faces were perceived to be sadder with sad music than with happy music. Across both conditions, neurotypical children rated the happy faces as happier and the sad faces as sadder than did participants with ASD. Response times of the neurotypical children were consistently shorter than response times of the children with ASD; both groups took longer to rate sad faces than happy faces. Response times of neurotypical children were generally unaffected by the valence of the music condition; however, children with ASD took longer to respond when listening to sad music.*

The author would like to thank Dr. Judith Jellison and Dr. Robert Duke for their assistance with this research study. This research was completed in partial fulfillment of a PhD in Music and Human Learning at the University of Texas at Austin.

Address correspondence concerning this article to Laura Brown, PhD, MT-BC, 1 Ohio University, Athens, Ohio, 45701. Phone: 740-593-4234. E-mail: BrownL5@ohio.edu

Conclusions: *Music appears to affect perceptions of emotion in children with ASD, and perceptions of sad facial expressions seem to be more affected by emotionally congruent background music than are perceptions of happy or neutral faces.*

Introduction

Although individuals with autism spectrum disorder (ASD) present with a range of behavioral characteristics, at the core of an ASD diagnosis are social and communication deficits (American Psychiatric Association, 2013). One common social problem experienced by individuals with ASD is emotion perception. In fact, Kanner's (1943) first description of ASD was "autistic disturbances of affective contact" (p. 217). Many individuals with ASD have difficulty perceiving emotions in faces, yet appear to be able to identify emotions in music. Researchers are currently exploring how, why, and to what degree emotion perception is compromised in individuals with ASD, as well as how these individuals perceive emotion in music.

Many children with ASD have difficulty identifying facial emotions (Harms, Martin, & Wallace, 2010). Individuals with ASD have different gaze patterns than do neurotypical peers when simply *viewing* facial expressions, without being asked to actually identify the emotion (Klin, Jones, Schultz, Volkmar, & Cohen, 2002). Children and adults with ASD often avoid gazing at the eyes and focus less on central facial features (i.e., eyes, nose, mouth) than do their neurotypical peers (Jones, Carr, & Klin, 2008; Pelphrey et al., 2002). It appears that adults with ASD use a feature-based strategy instead of the holistic process that is used by most neurotypical individuals (Hubl et al., 2003), meaning that individuals with ASD view faces in parts (e.g., eyes, then mouth, then nose) rather than as a whole.

Some research shows that children with ASD have a facial emotion recognition (FER) deficit (Celani, Battacchi, & Arcidiacono, 1999; Gross, 2004); other research has found no difference in FER accuracy, yet reveals other important differences in how neurotypical children and children with ASD process emotions. Children with ASD have demonstrated accurate FER when faces are presented slowly (Gepner, Deruelle, & Grynfeltt, 2001), and accuracy of FER does not appear to be affected when faces are inverted

(Rosset et al., 2008), showing that individuals with ASD can accurately interpret emotion when given enough time and further supporting the idea that faces are viewed as a sum of parts rather than a whole. Intelligence also appears to impact FER accuracy, as children with high-functioning ASD are better able to perceive emotions from faces than are children with both ASD and intellectual impairments (Dyck, Piek, Hay, Smith, & Hallmayer, 2006). Familiarity also influences FER accuracy; neurological responses to familiar faces tend to be similar in children with ASD and neurotypical peers (Pierce & Redcay, 2008). Non-emotional familiar stimuli, such as a mother's hand, have also been shown to elicit typical brain activity in children with ASD (Oberman, Ramachandran, & Pineda, 2008). The inconsistent findings related to FER accuracy warrant further study about how children with ASD recognize emotions.

Even though many individuals with ASD have difficulty perceiving emotion from faces, some individuals with ASD appear to be able to identify the emotional intent of musical excerpts. In an early pilot study, Heaton, Hermelin, and Pring (1999) found that children with ASD could perceive simple emotions (happiness and sadness) from four-measure melodies (major or minor) played on a keyboard. Children with ASD also appear to be able to perceive emotion from longer, intact music recordings, such as excerpts from large orchestral works (Heaton, Allen, Williams, Cummins, & Happé, 2008). Neurological activity while listening to music appears to be similar in individuals with ASD and their neurotypical peers. Caria, Venuti, and de Falco (2011) used fMRI to scan adults with Asperger's syndrome and neurotypical adults while they listened to happy or sad classical music excerpts and found that both groups had brain activity in regions associated with emotion processing during the listening tasks. Adults with ASD and their neurotypical peers also appear to have similar physiological responses while listening to emotional music (Allen, Davis, & Hill, 2013).

Factors such as intelligence, familiarity, and speed of stimuli presentation affect how indviduals with ASD perceive faces; several of these factors also influence the perception of emotional intent in music. Quintin, Bhatara, Poissant, Fombonne, and Levitin (2011) found that verbal intelligence was related to the success of identifying the emotional intent of music. Children with ASD were less accurate than neurotypical peers at matching orchestral excerpts to line drawings of faces, except when verbal IQ scores were factored out as a covariate (resulting in no difference between the two groups). Response time also tends to influence how individuals with ASD perceive emotion in music. DeBruyn, Moelants, and Leman (2012) found that adults with ASD were able to associate emotional music with words in a similar fashion as their neurotypical peers, though the group with ASD took longer to complete the task than did the neurotypical participants.

Given that some children with high-functioning ASD appear to be able to identify the emotional intent of music, yet struggle with perceiving facial emotion, exploring the relationship between these two modalities may inform possible interventions for enhancing emotional understanding in children with ASD. Humans quite naturally form associations among different sensory experiences, and signals obtained from different senses have been shown to influence one another. Music that connotes a happy feeling, for example, is rated happier if the performer looks happy while performing than if the performer displays a different emotion or is not visible to the listener (Thompson, Graham, & Russo, 2005; Thompson, Russo, & Quinto, 2008). Emotion perception of visual stimuli can also be primed by emotional music. For example, in a study by Logeswaran and Bhattacharya (2009), adults rated facial expressions as more happy when happy music was playing, more sad when sad music was playing, and neutral faces were rated in the direction of the musical prime.

Findings are mixed regarding multimodal emotion perception (e.g., listening to auditory stimuli while viewing visual stimuli) in individuals with ASD. Some studies reveal that individuals with ASD show a deficit when matching emotions from multimodal stimuli, such as matching an emotive speech example to a photograph of facial expressions (Hall, Szechtman, & Nahmias, 2003; Lindner & Rosén, 2006; Loveland, Steinberg, Pearson, Mansour, & Reddoch, 2008; Philip et al., 2010), while others show deficits for certain emotions, such as surprise (Jones et al., 2011) and fear (Magnée, de Gelder, van Engeland, & Kemner, 2008). Familiarity with the stimuli (e.g., pictures of self or family members) may also play a role in how well individuals with ASD can perceive emotions from multimodal stimuli (Kahana-Kalman & Goldman, 2008), and IQ may be a stronger predictor of the ability to perceive emotions from multimodal stimuli than diagnosis (Jones et al., 2011). All of the studies referenced above used recorded speech as the audio component of the multimodal emotion presentation. To date, music has not been investigated as the auditory source in multimodal emotion perception studies with children with ASD.

Emotional recognition is an important component of healthy social functioning in children with ASD and is an emerging area of focus in music therapy research (Geretsegger, Elefant, Mössler, & Gold, 2014, LaGasse, 2014; Whipple, 2004). A recent survey of music therapists working with children with ASD revealed that 43.1% of respondents were focusing on emotional goals in treatment with their clients (Kern, Rivera, Chandler, & Humpal, 2013). Music therapy studies have examined several factors related to emotions and emotional processing, including problem behaviors, communication, joint attention, and overall social functioning (Geretsegger, Elefant, Mössler, & Gold, 2014); however, to date only one music therapy study has directly examined the use of music therapy interventions to improve accuracy of labeling emotions and the expression of emotions in children with ASD (Katagiri, 2009). In this study, instruction with recorded improvised background music was more effective than instruction with singing songs, verbal instruction alone, and no intervention. Before and after the training procedure, participants were asked to label happy, sad, angry, and fearful emotions in faces from photographs, drawings, and schematic pictures, and to make different emotional facial expressions when prompted. Based on the findings, the author concluded that recorded improvised background music was the most successful intervention for improving recognizing and producing facial emotions. In addition, parent reports showed that participants enjoyed both the background music and singing conditions more than verbal instruction or no instruction.

Given the growing body of music therapy research that is focusing on children with ASD (Brown & Jellison, 2012), it seems important to examine more closely the relationships between the perception of emotion in facial expressions and the perception of emotion in music. Music has been shown to influence neurotypical adults' perceptions of realistic facial photographs (Logeswaran & Bhattacharya, 2009), yet there is little information about music's influence among children or adults with ASD engaged in this same task. To date there is no investigation of the influence of music with strong emotional valence on the ratings of facial photographs by children with ASD and their neurotypical peers. The purpose of this study was to investigate the influence of music on children's ratings of emotions expressed in human faces. Specifically, the following research questions were addressed:

1. Are children with ASD and neurotypical children differentially affected by music with strong emotional valence in their ratings of emotions in human faces?
2. Are children with ASD and neurotypical children differentially affected by music with strong emotional valence in their response times when rating emotions in human faces?

Method

Overview

To investigate the influence of music with strong emotional valence on children's recognition of emotions depicted in human faces, neurotypical children and children with high-functioning ASD listened to music intended to convey happiness and then to convey sadness as they identified and rated the emotions in 30 photographs of faces under both the happy and sad music conditions. I compared the ratings for the two participant groups as well as the latencies of their responses. The Institutional Review Board at the University of Texas at Austin approved the research protocol, and parent and guardian consent was obtained for all participants.

Participants

Participants were 20 children with ASD and 30 neurotypical (NT) children, all between the ages of 6 and 13 years old. Children with ASD were recruited from elementary and middle schools, and neurotypical children were recruited from an afterschool program at an elementary school and a middle school. I selected children with ASD based on the children's ability to complete the tasks required for the study and in consultation with their teachers, who were asked to recommend high-functioning students for participation. With parents' and guardians' permission, participants' Individual Education Plans (IEPs) were reviewed to confirm a diagnosis of ASD and to ensure that they were receiving special education services based on their diagnosis. High-functioning autism (HFA) is not a separate special education category, so IEPs did not indicate a separate diagnosis of HFA. Teacher expertise and the screening procedures described below were used to eliminate those participants that did not qualify to complete the study.

Only children who successfully completed the training procedure and attended to the task for the duration of the procedure were included in the study. Initially 24 participants were recruited for the ASD group, although four (all male) were unable to successfully complete the training protocol and were therefore not included in the study. Of the 20 children with ASD who were included, 18 were male and the mean age of this group was 9.10 years (SD = 1.82). Thirty neurotypical children participated; 18 were male, and the mean age of this group was 9.63 (SD = 1.73). The difference between the groups in the proportion of males to females was expected given the higher incidence of ASD in males than in females (Centers for Disease Control and Prevention, 2014).

Materials

The photographs used in this study were from a collection of 637 photos entitled the NimStim set (Tottenham et al., 2009). The set contains photographs of African American, European American, Latino American, and Asian American adult actors showing happy, sad, disgusted, fearful, angry, surprised, neutral, and calm faces. The NimStim photographs were originally validated by Tottenham et al. (2009), who examined responses of 81 adults who labeled the photographs from a choice of six emotion words.

The final set of photographs for the present study consisted of 30 photographs (10 each of happy, neutral, and sad) selected from the NimStim set. Photographs with moderate accuracy ratings (between 60% and 80%) by adults (Tottenham et al., 2009) were selected because visual stimuli that are ambiguous in nature have been shown to be influenced by auditory stimuli (Massaro & Egan, 1996). I arrived at a subset of 70 photographs with moderate consensus scores, and from this subset arrived at a final set of 30 photographs for this study (a list of photographs is available from the author).

The final set of 30 photographs with moderate consensus scores were selected by randomly ordering the photographs representing each of the three emotions, and then choosing the first five photographs of males and the first five photographs of females in each order with the restriction that actors would not be repeated. I also eliminated several photographs that may have appeared scary or startling to young children. The 10 photographs of faces in each of the three emotion categories were of actors of various ethnic backgrounds. Several photographs with rating scores outside the moderate range had to be included so that actors were not repeated and an equal proportion of each gender was maintained in each set. The mean accuracy percentage from Tottenham et al. (2009) for the 10 happy photographs was 0.812 (*SD* = 0.117). The happy photographs have higher mean accuracies than either the sad or neutral photos, since happy appears to be a more easily recognized emotion (Elfenbein & Ambady, 2002). Happy ratings in other photo sets also have higher accuracy ratings than other emotions, such as disgust or sadness (Ekman & Friesen, 1975). The mean accuracy percentage for the 10 sad photographs was 0.772 (*SD* = 0.122), and the mean accuracy percentage for the 10 neutral photographs was 0.691 (*SD* = 0.046).

Stimulus Presentation. The experimental stimuli were presented using SuperLab® software (Abboud, Heller, Matsak, Schultz, & Zeitlin, 2012), a RB-730 Response Pad, a MacBook Pro laptop, and an iPod fitted with external speakers. The software presented the photographs and recorded each participant's response and response latency. The response pad had seven keys (without numbers) and was outfitted with an icon indicating the continuum of very happy to very sad faces (see Figure 1). The iPod was used to play the music during the experiment.

Video Recordings. Participants were recorded using a Flip HD camera and the laptop camera on the Macbook Pro. The Flip camera was set up with the computer and response pad in view, and the laptop camera was focused with each participant's face in view.

Independent Variables

A stringent procedure was used to select four music stimuli for the study. Previous research investigating emotion perception in Western classical music was consulted for music recording sources (Logeswaran & Bhattacharya, 2009; Mitterschiffthaler, Fu, Dalton, Andrew, & Williams, 2007). After consultation with music professors, I used seven selections from Mitterschiffthaler and colleagues (2007) for further examination. The seven selections were played for college students (non-music majors), who labeled four emotions for each recording. Four of the seven recordings were consistently labeled as happy or sad.

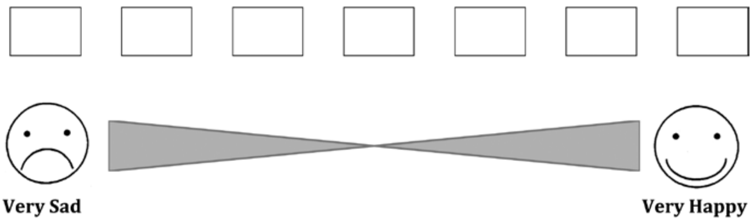

FIGURE 1.
Response pad icon.

As additional confirmation, the four selections were played for six neurotypical children (ages 6–13), who were asked to think of a happy and sad story and then select a happy and sad piece of music to go with that story. The two selections most often chosen by the children to represent happy and sad were "La Primavera (Spring): Allegro" by Vivaldi (Salzburg Chamber Orchestra, 1991) and "Adagio for Strings" by Barber (New York Philharmonic, 2004), respectively. Both recordings were looped to play on the iPod continuously for approximately 10 minutes. Only the first 5 minutes of the "Adagio for Strings" was looped because the second half of the piece abruptly changes emotional intent (e.g., tempo and modality changes).

Dependent Variables

There are two dependent variables in this study: ratings of facial photographs and latency of rating responses. Ratings corresponded with the seven-key response pad, and latency was measured in milliseconds from the onset of the photograph on the computer screen to the participants' key press. Responses for ratings and response latencies were collected using the SuperLab® software. I recorded participants' comments that were relevant to the study using paper and pencil.

Procedures

The procedures were identical for both the neurotypical participants and the children with ASD. Each participant was tested individually. The experiment took place at the schools of the participants in a quiet, isolated room with a table, two chairs, and the experimental equipment. Participants were seated in front of the laptop computer with the response pad placed in front of the laptop keyboard and the iPod placed to the side. The speakers were placed on either side of the computer, and I was seated to the right of the participant. The volume of the music was set at a listening level that appeared comfortable for the participant. If the participant exhibited any signs of discomfort or spoke about the volume level when the music began playing, I adjusted it to his or her comfort level.

Prior to beginning the training procedures, I engaged each participant in conversation in order to make him or her as comfortable as possible. I avoided any conversation topics that could possibly arouse intense emotion (e.g., testing at school).

Training Procedures. All participants had training regarding using the response pad and responding to happy and sad stimuli. Participants were told that they would hear some music, see some pictures of faces, and decide how happy or sad the faces were. The participants were then introduced to the seven-key response pad labeled with the continuum of very sad to very happy. After a demonstration, participants were asked to touch a button for "very happy/very sad" faces, and touch a button for a face that was not as "happy/sad." Participants practiced pressing buttons for more or less happy/sad faces until they understood the possible range of happy/sad choices, including the middle button that meant that some faces would be neither happy nor sad.

Practice Trials. Practice trials of 4 faces (2 males and 2 females; 2 happy and 2 sad) were presented using the SuperLab® software. Participants were asked to decide how happy or sad the faces were by pressing the appropriate key on the response pad. If participants appropriately utilized the different keys on the response pad, the procedure continued. If the participants used only the extreme keys (very happy and very sad), they were asked to repeat the practice experiment and see if they could use some of the other buttons to decide *how* happy or sad the faces were. The four participants who did not appear to understand how to use the scale to rate the faces after practicing the procedure twice did not complete the remainder of the experiment.

Experimental Trials. Following the practice procedure, participants were told that they would hear some music, see faces, and then decide how happy or sad some faces were, just as they did when they were practicing. Participants heard 1 minute of either the happy or sad music while viewing an icon (a drawing of an ear with the word "Listen") on the computer screen. The order of the music selections (happy or sad first) was randomly determined.

After 1 minute of listening, the music continued, the screen cleared, and the 30 faces selected from the NimStim set began to appear in random order. Following each response, a blank screen appeared for 4 seconds. Participants had a total of 12 seconds to respond to each face, but the photographs advanced to the next blank screen as soon as a response was made. If participants did not make a response within 7 seconds, I prompted them to "make your best guess." If participants still did not respond, they were prompted once more at the 10-second mark. If participants still did not respond after 12 seconds, the photograph automatically advanced, and no data were recorded. The music was stopped after the 30 faces were presented.

Participants then took a 3-minute break during which they had a choice of coloring a picture, completing a word search game, or doing nothing. No conversation occurred and no questions about the study or procedures were answered during the break. I gave participants a warning when there was 1 minute left in the break and let them know they could take their activity home if they desired.

After the break, participants were told that they would hear some more music and continue to decide how happy or sad the faces were. The same "listen" icon appeared on the screen, and the next music (either happy or sad) was played for 1 minute. I then reminded participants that the faces would soon appear and that they should decide how happy or sad the faces were. The same 30 faces were then presented (with a blank screen in between for 4 seconds) in a new random order. The same prompting procedures were used as in the first music condition.

A separate no-music condition for this study was considered during pilot testing, but the procedure was too long for the participants with ASD to attend to the task.

Statistical Analysis

Facial Ratings. Facial ratings were converted to standard scores by subtracting the grand mean for each participant's 60 ratings (30 faces rated during happy and sad music conditions) from each individual score and then dividing by the standard deviation of the scores for that participant. Standardized scores represent each participant's comparative ratings among the 60 presentations. I conducted a three-way mixed analysis of variance (ANOVA) with repeated measures using the 30 photographs as "subjects" (10 happy faces, 10 neutral faces, and 10 sad faces). Means of the standard scores of ratings for each photograph under the two music conditions resulted in a total of 120 data points (30 faces rated under happy and sad music by two groups of participants). In the three-way ANOVA, face type (happy, neutral, and sad) was the between factor and music type (happy and sad) and group (NT and ASD) were the within factors. This allowed for an equal sample size for the between factor, as the number of participants in the ASD and NT groups were unequal. Missing data points (ASD $n = 3$, NT $n = 1$) were replaced with averages of each participant's ratings under the same condition (e.g., happy face ratings while listening to happy music). Data were analyzed using SPSS software. I found unequal variances in participant ratings according to Levene's test; in this analysis, I used Greenhouse-Geisser adjusted degrees of freedom. I used Scheffe's method for post hoc analysis. Effect sizes were reported using partial eta-squared (ηp2).

Response Latency. I performed a repeated measures ANOVA to analyze differences in response latencies, again treating individual photographs as subjects and using the mean response latencies for each photograph under each condition as each photograph's score. Face type (happy, neutral, or sad) was the between factor, and music type (happy or sad) and group (ASD or NT) were the within factors in the analysis. Missing data points for response latency were treated in the same manner as missing data points for facial ratings. Here, too, I replaced missing data points with the averages for each participant for each condition. According to Levene's test, variances were equal for the response times. Both Kolomogorov-Smirno's and Shapiro-Wilks's tests indicated that the latency data were normally distributed ($p > .136$).

For both facial ratings and response latency, higher-order interactions were first considered for significance and were removed from the model if not significant.

Results

Ratings

I first determined whether participants rated the faces in the appropriate categories (happy, neutral, sad). Group means and standard deviations for response ratings 1 through 7 are reported in Table 1. The means in Table 1 show that participants in both groups generally used the scale of 1–7 to appropriately identify the intended emotion, although the responses by the participants with ASD were more varied than those of the neurotypical participants, as evidenced by larger standard deviations for all facial ratings. Standard scores of the ratings appear in Table 2. Neurotypical boys' and girls' ratings were compared to ensure that there was no difference between groups based on gender, $F(1,28) = 1.478$, $p = .23$.

Influence of Music on Ratings of Emotion

The analysis of the standard scores of facial ratings is presented in Table 3. There was a significant interaction between face type and music condition, $F(2,27) = 10.15$, $p = .001$, $\eta p2 = .43$, indicating that the influence of the two music conditions on standard scores was not consistent across face types (see Figure 2). The ratings of the happy and neutral faces were generally unaffected by the music conditions. Both happy and neutral faces were rated slightly

Table 1

Participant Facial Rating Means (standard deviations in parentheses)

	NT ($n = 30$)		ASD ($n = 20$)	
Face type	Happy music	Sad music	Happy music	Sad music
Happy	6.06 (0.95)	6.06 (1.05)	6.07 (1.21)	6.10 (1.26)
Neutral	3.76 (0.66)	3.73 (0.58)	3.54 (1.19)	3.62 (1.16)
Sad	2.32 (1.02)	2.19 (0.96)	2.36 (1.33)	2.09 (1.25)

Note. ASD = Participants with ASD; NT = neurotypical participants.

Table 2

Mean Participant Facial Rating Standard Scores (standard deviations in parentheses)

	NT ($n = 30$)		ASD ($n = 20$)	
Face type	Happy music	Sad music	Happy music	Sad music
Happy	1.13 (0.39)	1.15 (0.39)	1.05 (0.28)	1.07 (0.29)
Neutral	–0.15 (0.07)	–0.16 (0.08)	–0.19 (0.10)	–0.16 (0.12)
Sad	–0.94 (0.29)	–1.03 (0.28)	–0.81 (0.43)	–0.95 (0.33)

Note. ASD = Participants with ASD; NT = neurotypical participants.

TABLE 3

Results of Three-Way Analysis Variance of Group x Music x Face Type for Ratings

Source	SS	df	MS	F	p	η_p^2
Music (M)	.02	1	.02	4.60	.04	.15
Group (G)	.000	1	.000	.000	1.00	.000
Face type (FT)	84.54	2	42.27	148.76	.000	.92
MxG	.000	1	.000	.002	.97	.000
MxFT	.10	2	.05	10.15	.001	.43
GxFT	.17	2	.08	4.11	.03	.23
MxGxFT	.01	2	.01	.76	.48	.05

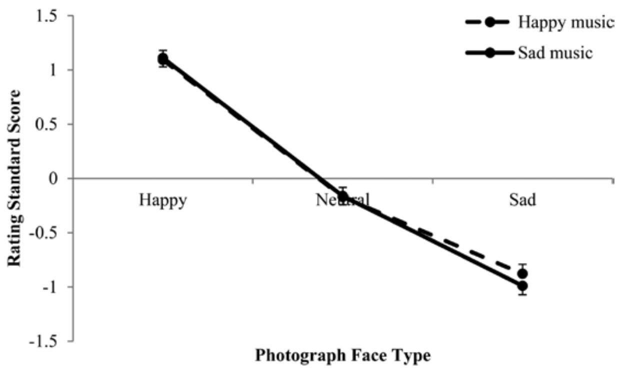

FIGURE 2.
Emotion ratings: Music by face type interaction, $F(2,27) = 10.15$, $p = .001$, $\eta_p^2 = .43$. Error bars represent the standard error of the mean (SEM).

higher (happier) during the sad music condition as compared to the happy music condition. In both groups, sad faces were rated as more sad under the sad music condition as compared to the happy music condition.

A significant group by face type interaction was found, $F(2,27) = 4.11$, $p < .03$, $\eta_p^2 = .23$, showing that the two groups rated the face types differently. Neurotypical participants rated happy and neutral faces higher (happier) than did participants with ASD, and neurotypical participants rated the sad faces lower (sadder) than did participants with ASD (see Figure 3). There was no significant interaction between music and group, $F(1,27) = 0.002$, $p > .05$, and no significant three-way interaction between music condition, group, and face type, $F(2,27) = 0.76$, $p > .05$. There was no significant main effect for group because standardized scores were used; all group means were zero.

Response Latency

Means and standard deviations for response latencies (in milliseconds) appear in Table 4. Results of the ANOVA are presented in Table 5. There was no significant music by face type interaction, $F(2,27) = .16$, $p > .05$, indicating that the music and face type factors were independent. There was a significant group by face type interaction, $F(2,27) = 4.23$, $p < .03$, $\eta_p^2 = .24$. Participants with ASD rated happy faces the fastest, followed by sad faces and then neutral faces. Neurotypical participants rated happy faces the fastest, followed by neutral faces and then sad faces (see Figure 4). Participants with ASD were also affected more by the music than neurotypical participants, as evidenced by a significant group by music interaction,

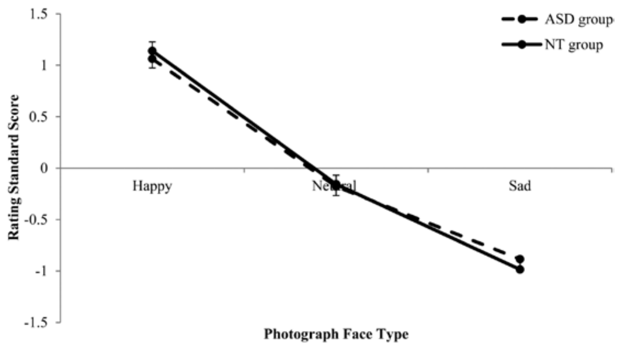

FIGURE 3.
Emotion ratings: Group by face type interaction, $F(2,27) = 4.11$, $p < .03$, $\eta_p^2 = .23$. Error bars represent SEM.

TABLE 4

Mean Participant Response Times in Milliseconds (standard deviations in parentheses)

	NT (*n* = 30)		ASD (*n* = 20)	
Face type	Happy music	Sad music	Happy music	Sad music
Happy	2159.78 (312.89)	2201.63 (345.37)	2594.41 (386.18)	2820.83 (418.99)
Neutral	2294.60 (212.61)	2286.17 (238.47)	2993.11 (369.97)	3251.06 (360.21)
Sad	2463.12 (212.54)	2561.22 (197.63)	3036.70 (329.73)	3113.92 (467.00)

Note. ASD = Participants with ASD; NT = neurotypical participants.

TABLE 5

Results of Three-Way Analysis Variance of Group x Music x Face Type for Response Latency

Source	SS	df	MS	F	p	η_p^2
Music (M)	400331.38	1	400331.34	10.85	.003	.29
Group (G)	12310314.10	1	12310314.10	187.80	.000	.87
Face type (FT)	2647215.56	2	1323607.78	4.34	.02	.24
MxG	154128.34	1	154128.34	4.67	.04	.15
MxFT	12078.74	2	6039.37	.16	.85	.01
GxFT	554439.41	2	277219.71	4.23	.03	.24
MxGxFT	109519.95	2	54759.56	1.657	.21	.11

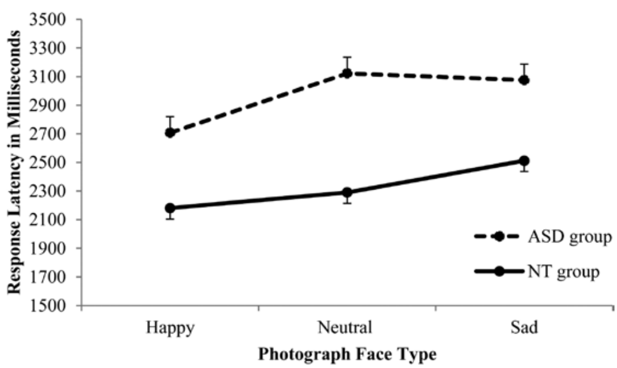

FIGURE 4.
Response latency: Group by face type interaction, $F(2,27) = 4.23$, $p < .03$, $\eta_p^2 = .24$. Error bars represent SEM.

$F(1,27) = 4.66$, $p = .04$, $\eta_p^2 = .15$ (see Figure 5). Participants with ASD took an average of 187 milliseconds longer to rate faces while listening to sad music than while listening to happy music, and neurotypical participants took an average of 43 milliseconds longer to rate faces while listening to sad music than while listening to happy music. There was no significant three-way interaction between face type, music condition, and group, $F(2,27) = 1.66$, $p > .05$.

Discussion

Overall Emotion Recognition

Group means revealed that both groups of participants, overall, rated the 30 faces correctly using the three emotional categories

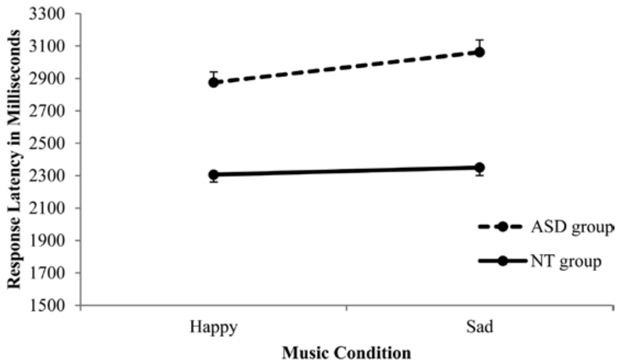

FIGURE 5.
Response latency: Group by music interaction, $F(1,27) = 4.66$, $p = .04$, $\eta_p^2 = .15$. Error bars represent SEM.

(happy, neutral, and sad). This result may be due to the high-functioning level of the participants with ASD in this study. Many children with ASD struggle with facial emotion recognition, but higher-functioning children can have similar responses to neurotypical peers under certain conditions, such as having familiar stimuli or more time to complete the task (Harms et al., 2010). In fact, intelligence and emotional functioning are more strongly correlated for children with ASD than for neurotypical children (Dyck, Piek, Hay, Smith, & Hallmayer, 2006). Neurotypical individuals often have automatic neurological responses to emotions, whereas individuals with ASD often process emotional faces as if they were objects (Schultz et al., 2000). This object-like processing can result in accurate facial emotion recognition, though it often takes longer (Hubl et al., 2003). The current investigation also found that response time was slightly longer for children with ASD than it was for neurotypical peers. This finding will be discussed later in greater detail.

Although neurotypical participants and participants with ASD identified faces in three emotion categories accurately, participants with ASD had greater variation in their ratings than did the neurotypical group, as evidenced by larger standard deviations (see Table 1). The standard deviation for ratings of neutral faces for the ASD group was .53 larger for the happy music condition and twice as large under the sad music condition. The increased variance in responses might be explained by a true misunderstanding of the faces, influence of the music, or simply a lack of focus. Even though participants with ASD were high functioning, video recordings during the sessions revealed that several had short lapses in attention (e.g., looking away and requiring a prompt to focus). Neurotypical participants required almost no prompts to remain focused; several participants with ASD required repeated prompting to stay on task.

Influence of Music on Facial Ratings

Both groups of participants were influenced by listening to music when rating sad faces; ratings for sad faces were lower in the sad music condition compared to the happy music condition. Surprisingly, there was little influence of the music (happy or sad) on ratings of happy or neutral faces. In fact, both groups rated the happy and neutral faces as slightly happier during the sad music condition than during the happy music condition, though these differences were small and nonsignificant.

The fact that the two pieces of music influenced both groups' ratings of emotional faces provides additional evidence that children with ASD and neurotypical children respond similarly to emotion in music. There is recent evidence (using music other than the two pieces used in this study) that individuals with ASD and neurotypical peers show similar brain activity when listening to music with a strong emotional valence (Caria, Venuti, & de Falco, 2011), and that both groups can accurately label emotional intent of music excerpts by pointing to facial drawings (Heaton, Hermelin, & Pring, 1999; Heaton, Allen, Williams, Cummins & Happé, 2008; Quintin, Bhatara, Poissant, Fombonne, & Levitin, 2011).

Responses were similar for the two groups, but what remains unanswered is why only the ratings for sad faces were influenced by the music. One possibility is that happy facial expressions are more easily recognized than are neutral and sad expressions, and thus may be less influenced by music. Recall that in the NimStim photograph set, the original validation procedures showed that there was a greater consensus for identifying the happy photographs as compared to the consensus ratings for other emotions (Tottenham et al., 2009). Not surprisingly, then, the happy faces chosen from the set for the current study also had greater consensus ratings than had either the sad or neutral faces: the mean rating for happy faces was 0.812 (SD = 0.117), the mean rating for sad faces was

0.772 (*SD* = 0.122), and the mean rating for neutral faces was 0.691 (*SD* = 0.046). Happy faces in other photograph sets used for emotion recognition also have higher accuracy ratings than do photographs of other types of emotional faces (Ekman & Friesen, 1975).

Several studies suggest that children with ASD rate happy faces more accurately than they rate faces depicting other emotions. Castelli (2005) found that in an array of angry, sad, fearful, disgusted, and surprised faces, children with ASD rated happy faces most accurately. Katagiri (2009) found that children with ASD showed little improvement in the recognition of happy faces after participating in music sessions focused on emotion recognition training; the children performed well on pretest measures for identifying happiness at the start of the study, and there was little room for improvement. There was improvement, however, from pretest to posttest on tasks for identifying anger, sadness, and fear.

Elfenbein and Ambady (2002) found happiness to be the most accurately recognized emotion in their meta-analysis of emotion recognition in and across cultures. The meta-analysis also found that happiness was the easiest emotion to recognize from faces and the hardest to recognize from voices. In contrast, sad emotions were some of the easiest emotions to recognize from voices, yet more difficult to recognize from faces. Perhaps the emotion of sadness is more saliently conveyed through auditory forms rather than visual forms. In their empirical review of emotion research and speech, Juslin and Laukka (2003) found that consensus of emotion identification was similar for both speech and music, and that sadness and anger were the two emotions most clearly conveyed through these auditory sources, as compared to fear, happiness, and tenderness. If there is less consensus about what constitutes a sad face than a happy face, perhaps sad faces may be more readily influenced by other stimuli, such as the music in this study, particularly if that stimuli clearly conveys a sad emotion through an auditory source.

A surprising finding from the current study was the apparent lack of influence by music on ratings of the neutral photographs. Previous research shows that ratings of neutral faces are more influenced by background music than are happy and sad faces (Logeswaran & Bhattacharya, 2009). The music selections in the current study were different than those used in Logeswaran and Bhattacharya, and although the photograph set was the same, I was unable to determine the exact photographs used by Logeswaran and Bhattacharya. It is possible that differences in stimulus materials explain the differences in results. The music excerpts used in Logeswaran and Bhattacharya were drawn from a study investigating positive and negative valence (Altenmüller, Schürmann, Lim, & Parlitz, 2002), not happy and sad emotions. The excerpts that they labeled as positive and negative were not necessarily representative of happiness and sadness. Additionally, the participants in Logeswaran and Bhattacharya were adults, who may be more influenced by contextual cues like music when evaluating the emotional expressions in faces.

Response Latency

Mean response times across all photographs were longer for the children with ASD than for the neurotypical children. Response times were longer during the sad music than during the happy music; this effect was more pronounced in participants with ASD than neurotypical participants. There were differences between the two groups in response times for each face type. Out of happy, neutral, and sad faces, participants with ASD took longest to rate neutral faces whereas neurotypical participants took longest to rate sad faces.

Several studies of neurotypical children and high-functioning children with ASD show that groups may appear similar on emotion recognition behavioral tasks, but slight differences have been found in neural activity (Hubl et al., 2003), eye gaze (Pelphrey et al., 2002), or time to complete a task (Gepner, Deruelle, & Grynfeltt, 2001). Slower response times in the current study may be evidence that participants with ASD need slightly more time to reach a decision compared to neurotypical participants. If individuals with ASD are using cognitive strategies instead of a more automatic neurological process to assess faces, perhaps even the slightest increase in viewing time helps them perceive the facial emotion more accurately.

As noted earlier, there is a strong consensus about what constitutes a happy face, as evidenced by the high proportions of adults who identified the happy faces in the NimStim set and other photograph arrays (Ekman & Friesen, 1975). It is possible that because happiness is an easily recognized emotion, responses in both groups were quicker for those faces than those for either sad or neutral faces.

The response latencies of the neurotypical participants and the participants with ASD appeared to have been influenced by the music; both groups combined were slower to rate faces when listening to sad music than when listening to happy music. Mean response times for both groups combined were 115 milliseconds slower while listening to sad music than while listening to happy music.

The response time group by face type interaction revealed that out of the three face types, participants with ASD took longest to rate neutral faces and neurotypical participants took longest to rate sad faces. Perhaps the increased response times by children with ASD for neutral faces shows that children with ASD are less clear about what constitutes a neutral face. Happiness and sadness are clear opposites, whereas neutral is a more ambiguous concept. Additionally, happiness and sadness are likely the first emotions to be taught to children, therefore the concepts are more familiar.

Recall that there was no interaction between group and music type for the mean ratings of emotion. For latency, however, a significant interaction was found for group and music type. The difference between the response latencies of the two groups was smaller under the happy music condition (569 milliseconds) than under the sad music condition (712 milliseconds). While listening to sad music, participants with ASD were slower at rating faces than were neurotypical participants. Future studies are needed to examine this observed difference in greater detail, and to determine whether response times of children with ASD are more affected by the music than the response times of neurotypical peers.

Limitations

Something that needs to be considered as a potential limitation to this project is the method of analysis. I analyzed the data as if each photograph were a subject, whose score was the mean rating by each group in each music condition, thus the results reflect the variance among photographs and not the variance among actual participants. In addition, by using this approach the standard errors may be underestimated. As such, I have worked to exercise caution in my presentation and interpretation of findings. Readers should also take caution in interpreting significant findings, as inferences are being made about individuals based on group-level data.

One challenge faced by many researchers working with children with ASD is finding an adequate sample size of participants with similar characteristics. In this study, the participants' diagnosis of ASD was confirmed via IEP, but the only measure of HFA was teacher report, observation by the researcher, and the ability to complete the training protocol. In future studies, the use of a larger sample of participants and the use of objective measures to confirm the diagnosis of HFA would strengthen the validity of study findings.

Conclusions

To my knowledge, this study is the first to establish the influence of music on facial emotion recognition in children with ASD and provides preliminary evidence that music may affect facial emotion recognition in children with ASD. In this study, sad faces were perceived to be sadder when participants were listening to sad music versus happy music. Consistent with previous studies, response times of children with ASD were longer than those of neurotypical children, and the response times of children with ASD were more affected by the music than those of neurotypical children. This preliminary evidence is important to consider in the development and interpretation of future research concerning ways children with ASD perceive, learn about, and express emotions, and in the development of future music therapy interventions.

References

Abboud, H., Heller, K., Matsak, E., Schultz, H., & Zeitlin, V. (2012). *SuperLab*. San Pedro, CA: Cedrus Corporation.

Allen, R., Davis, R., & Hill, E. (2013). The effects of autism and alexirhythmia on physiological and verbal responsiveness to music. *Journal of Autism and Developmental Disorders*, *43*(2), 432–444. doi:10.1007/s10803-012-1587-8

Altenmüller, E., Schürmann, K., Lim, V. K., & Parlitz, D. (2002). Hits to the left, flops to the right: Different emotions during listening to music are reflected in cortical lateralisation patterns. *Neuropsychologia*, *40*(13), 2242–2256. doi:10.1016/S0028-3932(02)00107-0

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Brown, L. S., & Jellison, J. A. (2012). Music research with children and youth with disabilities and typically developing peers: A systematic review. *Journal of Music Therapy*, *49*(3), 335–364. doi:10.1093/jmt/49.3.335

Caria, A., Venuti, P., & de Falco, S. (2011). Functional and dysfunctional brain circuits underlying emotional processing of music in autism spectrum disorders. *Cerebral Cortex*, *21*(12), 2838–2849. doi:10.1093/cercor/bhr084

Castelli, F. (2005). Understanding emotions from standardized facial expressions in autism and normal development. *Autism*, *9*(4), 428–449. doi:10.1177/1362361305056082

Celani, G., Battacchi, M. W., & Arcidiacono, L. (1999). The understanding of the emotional meaning of facial expressions in people with autism. *Journal of Autism and Developmental Disorders*, *29*(1), 57–66. doi:10.1023/A:1025970600181

Centers for Disease Control and Prevention. (2014). Prevalence of autism spectrum disorders among children aged 8 years. Autism and Developmental Disabilities Monitoring Network, 11 sites, United States, 2010. *Morbidity and Mortality Weekly Report*, *63*(SS-02), 1–21.

DeBruyn, L., Moelants, D., & Leman, M. (2012). An embodied approach to testing musical empathy in participants with an autism spectrum disorder. *Music and Medicine*, *4*(1), 28–36. doi:10.1177/1943862111415116

Dyck, M. J., Piek, J. P., Hay, D., Smith, L., & Hallmayer, J. (2006). Are abilities abnormally interdependent in children with autism? *Journal of Clinical Child and Adolescent Psychology*, *35*(1), 20–33. doi:10.1207/s15374424jccp3501_3

Ekman, P., & Friesen, W. V. (1975). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, *128*(2), 203–235. doi:10.1037/0033-2909.128.2.203

Gepner, B., Deruelle, C., & Grynfeltt, S. (2001). Motion and emotion: A novel approach to the study of face processing by young autistic children. *Journal of Autism and Developmental Disorders*, *31*(1), 37–45. doi:10.1023/A:1005609629218

Geretsegger, M., Elefant, C., Mössler, K. A., & Gold, C. (2014). Music therapy for people with autism spectrum disorder. *Cochrane Database of Systematic Reviews*, *2014*(6), 1–63. doi:10.1002/14651858.CD004381.pub3.

Gross, T. F. (2004). The perception of four basic emotions in human and nonhuman faces by children with autism and other developmental disabilities. *Journal of Abnormal Child Psychology: An Official Publication of the International Society for Research in Child and Adolescent Psychopathology*, *32*(5), 469–480. doi:10.1023/B:JACP.0000037777.17698.01

Hall, G., Szechtman, H., & Nahmias, C. (2003). Enhanced salience and emotion recognition in autism: A PET study. *American Journal of Psychiatry*, *160*(8), 1439–1441. doi:10.1176/appi.ajp.160.8.1439

Harms, M. B., Martin, A., & Wallace, G. L. (2010). Facial emotion recognition in autism spectrum disorders: A review of behavioral and neuroimaging studies. *Neuropsychology Review*, *20*(3), 290–322. doi:10.1007/s11065-010-9138-6

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*(2), 171–182. doi: 10.1348/026151007X206776

Heaton, P., Hermelin, B., & Pring, L. (1999). Can children with autistic spectrum disorders perceive affect in music? An experimental investigation. *Psychological Medicine*, *29*(6), 1405–1410. doi:10.1017/S0033291799001221

Hubl, D., Bölte, S., Feineis-Matthews, S., Lanfermann, H., Federspiel, A., Strik, W., Poustka, F., et al. (2003). Functional imbalance of visual pathways indicates alternative face processing strategies in autism. *Neurology*, *61*(9), 1232–1237. doi:10.1212/01.WNL.0000091862.22033.1A

Jones, C. R. G., Pickles, A., Falcaro, M., Marsden, A. J. S., Happé, F., Scott, S. K., Sauter, D., et al. (2011). A multimodal approach to emotion recognition ability in autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, *52*(3), 275–285. doi:10.1111/j.1469-7610.2010.02328.x

Jones, W., Carr, K., & Klin, A. (2008). Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. *Archives of General Psychiatry*, *65*(8), 946–954. doi:10.1001/archpsyc.65.8.946

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

Kahana-Kalman, R., & Goldman, S. (2008). Intermodal matching of emotional expressions in young children with autism. *Research in Autism Spectrum Disorders*, *2*(2), 301–310. doi:10.1016/j.rasd.2007.07.004

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, *2*, 217–250.

Katagiri, J. (2009). The effect of background music and song texts on the emotional understanding of children with autism. *Journal of Music Therapy*, *46*(1), 15–31. doi:10.1093/jmt/46.1.15

Kern, P., Rivera, N., Chandler, A., & Humpal, M. (2013). Music therapy services for individuals with autism spectrum disorder: A survey of clinical practices and training needs. *Journal of Music Therapy*, *50*(4), 274–303. doi:10.1093/jmt/50.4.274

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry*, *59*(9), 809–816. doi:10.1001/archpsyc.59.9.809

LaGasse, A. B. (2014). Effects of a music therapy group intervention on enhancing social skills in children with autism. *Journal of Music Therapy*, *51*(3), 250–275. doi:10.1093/jmt/thu012

Lindner, J. L., & Rosén, L. A. (2006). Decoding of emotion through facial expression, prosody and verbal content in children and adolescents with Asperger's syndrome. *Journal of Autism and Developmental Disorders*, *36*(6), 769–777. doi:10.1007/s10803-006-0105-2

Logeswaran, N., & Bhattacharya, J. (2009). Crossmodal transfer of emotion by music. *Neuroscience Letters*, *455*(2), 129–133. doi:10.1016/j.neulet.2009.03.044

Loveland, K. A., Steinberg, J. L., Pearson, D. A., Mansour, R., & Reddoch, S. (2008). Judgments of auditory-visual affective congruence in adolescents with and without autism: A pilot study of a new task using fMRI. *Perceptual and Motor Skills*, *107*(2), 557–575. doi:10.2466/PMS.107.6.557-575

Magnée, M. J. C. M., de Gelder, B., van Engeland, H., & Kemner, C. (2008). Atypical processing of fearful face-voice pairs in pervasive developmental disorder: An ERP study. *Clinical Neurophysiology*, *119*(9), 2004–2010. doi:10.1016/j.clinph.2008.05.005

Massaro, D. W., & Egan, P. B. (1996). Perceiving affect from the voice and the face. *Psychonomic Bulletin & Review*, *3*(2), 215–221. doi:10.3758/BF03212421

Mitterschiffthaler, M. T., Fu, C. H. Y., Dalton, J. A., Andrew, C. M., & Williams, S. C. R. (2007). A functional MRI study of happy and sad affective states induced by classical music. *Human Brain Mapping*, *28*(11), 1150–1162. doi:10.1002/hbm.20337

New York Philharmonic. (2004). Adagio for strings. *On Romantic favorites for strings [CD]*. Sony.

Oberman, L. M., Ramachandran, V. S., & Pineda, J. A. (2008). Modulation of mu suppression in children with autism spectrum disorders in response to familiar or unfamiliar stimuli: The mirror neuron hypothesis. *Neuropsychologia*, *46*(5), 1558–1565. doi:10.1016/j.neuropsychologia.2008.01.010

Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B. D., & Piven, J. (2002). Visual scanning of faces in autism. *Journal of Autism and Developmental Disorders*, *32*(4), 249–261. doi:10.1023/A:1016374617369

Philip, R. C. M., Whalley, H. C., Stanfield, A. C., Sprengelmeyer, R., Santos, I. M., Young, A. W., Atkinson, A. P., et al. (2010). Deficits in facial, body movement and vocal emotional processing in autism spectrum disorders. *Psychological Medicine: A Journal of Research in Psychiatry and the Allied Sciences*, *40*(11), 1919–1929. doi:10.1017/S0033291709992364

Pierce, K., & Redcay, E. (2008). Fusiform function in children with an autism spectrum disorder is a matter of "who." *Biological Psychiatry*, *64*(7), 552–560. doi:10.1016/j.biopsych.2008.05.013

Quintin, E.-M., Bhatara, A., Poissant, H., Fombonne, E., & Levitin, D. J. (2011). Emotion perception in music in high-functioning adolescents with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *41*(9), 1240–1255. doi:10.1007/s10803-010-1146-0

Rosset, D. B., Rondan, C., Da Fonseca, D., Santos, A., Assouline, B., & Deruelle, C. (2008). Typical emotion processing for cartoon but not for real faces in children with autistic spectrum disorders. *Journal of Autism and Developmental Disorders*, *38*(5), 919–925. doi:10.1007/s10803-007-0465-2

Salzburg Chamber Orchestra. (1991). La primavera (spring): Allegro. *On Treasury of light classics* [CD]. Maestro.

Schultz, R. T., Gauthier, I., Klin, A., Fulbright, R. K., Anderson, A. W., Volkmar, F., Skudlarski, P., et al. (2000). Abnormal ventral temporal cortical activity during face discrimination among individuals with autism and Asperger syndrome. *Archives of General Psychiatry*, *57*(4), 331–340. doi:10.1001/archpsyc.57.4.331

Thompson, W. F., Graham, P., & Russo, F. A. (2005). Seeing music performance: Visual influences on perception and experience. *Semiotica, 156*, 203–227. doi:10.1515/semi.2005.2005.156.203

Thompson, W. F., Russo, F. A., & Quinto, L. (2008). Audio-visual integration of emotional cues in song. *Cognition and Emotion*, *22*(8), 1457–1470. doi:10.1080/02699930701813974

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., et al. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–249. doi:10.1016/j.psychres.2008.05.006

Whipple, J. (2004). Music in intervention for children and adolescents with autism: A meta-analysis. *Journal of Music Therapy*, *41*(2), 90–106. doi:10.1093/jmt/41.2.90