Journal of Autism and Developmental Disorders
https://doi.org/10.1007/s10803-020-04781-0

ORIGINAL PAPER

Effects of Emotional Music on Facial Emotion Recognition in Children with Autism Spectrum Disorder (ASD)

Gary L. Wagener[1] · Madeleine Berning[2] · Andreia P. Costa[1] · Georges Steffgen[1] · André Melzer[1]

Accepted: 4 November 2020

Abstract
Impaired facial emotion recognition in children with Autism Spectrum Disorder (ASD) is in contrast to their intact emotional music recognition. This study tested whether emotion congruent music enhances facial emotion recognition. Accuracy and reaction times were assessed for 19 children with ASD and 31 controls in a recognition task with angry, happy, or sad faces. Stimuli were shown with either emotionally congruent or incongruent music or no music. Although children with ASD had higher reaction times than controls, accuracy only differed when incongruent or no music was played, indicating that congruent emotional music can boost facial emotion recognition in children with ASD. Emotion congruent music may support emotion recognition in children with ASD, and thus may improve their social skills.

Keywords Autism spectrum disorder · Children · Emotion recognition · Music · Facial recognition task

> "Music is the shorthand of emotion. Emotions, which let themselves be described in words with such difficulty, are directly conveyed to man in music, and in that is its power and significance" – Leo Tolstoy, letter to Sophia Andreevna, January 16, 1905

Children with ASD often show signs of processing difficulties of affective stimuli and social cues, which leads to problems recognizing emotions in social interactions. This was confirmed in a number of studies that showed differences between individuals with ASD and neurotypical individuals (e.g., Lozier, Vanmeter, & Marsh, 2014; Uljarevic & Hamilton, 2013). In a formal meta-analysis with data from 48 papers and over 980 participants with ASD, Uljarevic and Hamilton (2013) concluded that there is emotion recognition difficulty in ASD and that there are also differences in recognition accuracy for different emotions in people with ASD. However, it must be noted that emotion recognition difficulty might be dependent on the depicted emotion. Uljarevic and Hamilton (2013) found that the recognition of happiness was only marginally impaired in individuals with ASD. The findings on impaired emotion recognition for people with ASD were corroborated in the meta-analysis of Lozier, Vanmeter and Marsh (2014). The authors looked at 43 studies and a total sample of 1545 participants (791 participants with ASD, 754 neurotypical participants) with regard to emotion recognition impairments in individuals with ASD, differences in recognition across different emotional expressions, and moderating factors like age or cognitive impairment. The authors concluded that individuals with ASD show significant emotion recognition impairments across different emotional expressions that worsen with age (Lozier et al. 2014). Rump, Giovannelli, Minshew, and Strauss (2009) showed in their study on the development of emotion recognition in individuals with ASD that emotion recognition was impaired across several age groups (8–12 years; 13–17 years; adults). In line with this, Kuusikko et al. (2009) found that emotion recognition was worse for children with ASD (≤ 12 years; $M = 10.9$) than for adolescents with ASD (≥ 12 years; $M = 14.3$). Individuals with ASD also need significantly more time to recognize emotions in facial expressions than neurotypical individuals (Bal et al. 2010; Nacewicz, Dalton Johnstone, Long, & McAuliff, 2006). However, it must be noted that some studies did not find differences between neurotypical individuals and individuals with ASD on emotion recognition time (Fink, de Rosnay, Wierda, Koot, & Begeer, 2014). For example, Tracy, Robins, Schriber, and Solomon (2011) found similar

✉ Gary L. Wagener
Lee.Wagener@gmail.com

1 Department of Behavioural and Cognitive Sciences, University of Luxembourg, 11, Porte des Sciences, 4366 Esch-sur-Alzette, Luxembourg

2 Institute of Psychology, University of Trier, Universitätsring 15, 54286 Trier, Germany

Published online: 17 November 2020

levels of performance in emotion recognition for neurotypical individuals and individuals with ASD. Nonetheless, not being able to process social emotional cues correctly has several impacts on the behavior of individuals with ASD. For example, the deficits in emotion recognition and understanding other people's feelings may ultimately result in the inability to respond appropriately to social cues, limiting the capabilities for social interaction (Baron-Cohen 1995).

Even though children with ASD show problems with identifying emotional faces (Harms, Martin, & Wallace, 2010), processing affect in musical cues seems intact (Heaton, 2009; Heaton, Hermelin, & Pring, 1999) even for a wide range of emotions such as anger, fear, love, triumph, and contemplation (Heaton, Hermelin, & Pring, 2008). Similarly, a fMRI study provided evidence that there is no difference between adults with ASD and neurotypical adults for the processing of emotional (happy and sad) music (Caria, Venuti, & de Falco, 2011). Furthermore, neurotypical adults and adults with ASD do not differ in their physiological response to music (Allen, Davis, & Hill, 2013) and they engage similar neural networks during processing of emotional music (Gebauer, Skewes, Westphael, Heaton, & Vuust, 2014; Molnar-Szakacs & Heaton, 2012). However, in one study it was found that compared to a control group, children with ASD had decreased skin conductance response in relation to music (Stephenson, Quintin, & South, 2016a), and it is debated whether physiological effects of music-evoked emotions may change from childhood to adulthood in ASD (Stephenson, Quintin, & South, 2016b). Nevertheless, this suggests that unlike the processing of emotional facial expressions and social interactions, the processing of (emotional) music might not be impaired in people with ASD.

For neurotypical adults, Logeswaran and Bhattacharya (2009) showed that the processing of emotional faces (neutral, happy, and sad) was influenced when congruent emotional music (happy and sad) was played beforehand. In this case, subsequent happy (or sad) facial expressions were also rated as happier (or sadder). An earlier study tested the effect of music on emotion labeling accuracy and emotional understanding of children with ASD (Katagiri, 2009). Twelve students with ASD received emotion recognition treatment over eight sessions. Results showed that playing background congruent music significantly improved participant's emotional understanding after eight individual treatment sessions (Katagiri, 2009). Apparently, emotional music can improve the emotion processing of children with ASD. In fact, interventions designed around music therapy have several benefits for individuals with ASD like, for example, an improvement in mood, behavior, language, sensory perception, and social skills (Shi Lin & Xie, 2016). In a meta-analysis of 10 randomized-controlled trials, music therapy was concluded to be effective and superior to placebo therapy or standard care in improving social skills (i.e. social interaction, non-verbal and verbal communicative skills, initiating behavior, and social-emotional reciprocity) in people with ASD (Geretsegger, Elefant, Mössler, & Gold, 2014). Furthermore, Brown (2017) tested the effects of congruent versus incongruent emotional (happy and sad) music on facial emotion ratings (happy, sad, and neutral) with neurotypical children and children with high-functioning ASD aged 6–13 years. Brown (2017) found that when sad music was played, sad faces were perceived as sadder across both neurotypical children and children with ASD. However, in that study it was only tested if the emotion ratings for emotional faces would be affected by emotional music, not if accompanying music had an effect on emotion recognition accuracy. Nonetheless, this points to the value of emotional music as a potential facilitating agent in emotion recognition for children with ASD.

The aim of the present study was to test whether children without ASD and children with ASD differ in their recognition accuracy of emotions in facial cues and in their reaction times while either congruent or incongruent emotional music was played. It was hypothesized that no music or incongruent music causes individuals with ASD to have significantly lower emotion recognition accuracy and significantly greater reaction times in a facial emotion recognition task than individuals without ASD. Similarly, it was hypothesized that autistic traits (i.e., higher score on the autism spectrum quotient) would be positively related to more errors in emotion recognition and higher reaction times. This relationship was believed to hold only for the incongruent and no music conditions. Furthermore, unlike Brown (2017), who included only children with high-functioning ASD, a range of different ASD sub-types (i.e. early infantile autism, atypical autism and Asperger's Syndrome) were included in this study. Furthermore, the validated short form of the autism spectrum quotient for children (Allison, Auyeung, & Baron-Cohen, 2012), as a highly predictive measure for ASD, was hypothesized to be associated with diminished emotion recognition accuracy and greater reaction times only in trials with incongruent and no music, as congruent music was hypothesized to make up for the emotion recognition difficulties.

Methods

Participants

Fifty children aged 8–12 (M = 9.94; SD = 1.32) participated in the present study. Nineteen of these children had a diagnosis of ASD, with different types: n = 5 children were diagnosed with early infantile autism, n = 8 children with Asperger's Syndrome, and n = 6 with atypical autism.

Table 1 Distribution of age, AQ-10 score, and score on the emotion recognition difficulties scale across groups with tests for group differences

Measure	Control children ($n = 31$)				Children with ASD				Group Differences		
	M	SD	Min	Max	M	SD	Min	Max	Test	*p*	*d*
Age	10.10	1.16	8	12	9.68	1.53	8	12	$F = 1.16$	0.29	0.32
AQ-10 child	1.61	1.50	0	6	6.89	2.33	2	10	**$U = 22$**	**<0.001**	n.a.
Emotion recognition difficulties	0.76	0.68	0	2.7	2.13	0.65	0.4	2.9	**$F = 49.72$**	**<0.001**	**2.05**

Elements in bold mark significant results on the 0.001 significance level

Age distribution across groups can be seen in Table 1. Nineteen children were female (38%) and 31 were male (62%). Of the 19 children with ASD, 7 (36.8%) were female and 12 (63.2%) were male. There was no significant difference for gender distribution across both conditions, $\chi^2 = 0.02$, $p = 0.895$. Only children without any auditory or visual impairments were included in the study. The children were recruited through direct contact with therapy centers and schools in XX and YY (to be indicated after peer review).

Procedure

The present study was conducted in the field, meaning children were tested either at home, at school, or at an intervention center. To limit the influence of environmental confounds it was made sure that the testing environment was calm and had sufficient lighting.

At the beginning of the study, parents or caregivers signed an informed consent sheet for their children. Before the beginning of the tasks, children had the opportunity to get to know the researcher and needs such as toilet visits, hunger, and thirst were met. The study was divided into two independent sections. Parents or caregivers filled out the AQ-10 as well as the questionnaires on demographics for the children, children's emotion recognition difficulties, and the relevance of music for their children. At the same time, children took part in a facial emotion recognition task. After the practice run, a short break of one to two minutes followed. After task completion, children were asked how they perceived the task, if they found it difficult, and whether they encountered problems during the task. There were nine trials in total, accounting to each of the three emotions (joy, sadness, and anger) with either congruent music, incongruent music, or no music played in the background. All participants (participants with ASD and participants without ASD) rated the same stimuli in the same order. At the end of the study, participants were thanked and given chocolate as reward.

Questionnaire Material

Demographics

In the demographics part of the questionnaire, parents and caregivers who filled out the questionnaires for the children were asked about the age and gender of the child. Then, they provided information about the ASD diagnosis of the child (only for those in the ASD group).[1] Furthermore, they indicated the everyday contact the child has with music, such as whether the child actively listens to music or plays a music instrument.

Autism Spectrum Quotient

The short form of the Autism Spectrum Quotient questionnaire for children (AQ-10 Child) was used (Allison et al., 2012). The parent-report questionnaire assesses the potential neurodevelopment impairment with 10 items rated on a 4-point-Likert scale ranging from 1 (*I strongly disagree)* to 4 (*I strongly agree*). The scale had very good internal consistency with a Cronbach's $\alpha = 0.90$. The scale serves as a "red flag" tool in the assessment of ASD and can be divided into five subscales: *attention to detail*, *attention switching*, *communication*, *imagination*, and *social* (e.g., "*I find it difficult to work out people's intentions*"). The AQ-10 child was used instead of the original autism spectrum quotient (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001) as the AQ-10 child was more economic with only ten items and there is little performance difference between the AQ-10 and the full AQ-50 (Booth et al., 2013). For analysis, items are coded with 0 for the ratings *I strongly disagree* or *I slightly disagree*, and 1 for the ratings *I slightly agree* and *I strongly agree*. Then, the sum is computed for each participant to attain a score between 0 (minimum) and 10 (maximum).

[1] The ASD diagnoses were coded using DSM-IV criteria. So, the children had also been diagnosed with different subtypes of ASD, like Asperger's syndrome. Therefore, in the results section, in addition to results on the ASD group, analyses on differences between the different subtypes are provided as additional information and indicator for future studies.

The cut-off of the scale is set at a score of 6 (Allison et al., 2012), meaning that a value of 6 or higher typically warrants suspicion for ASD and is cause for further inspection. The AQ-10 child has a reported sensitivity of 0.95 and a specificity of 0.97 (Allison et al., 2012). For the 31 control children, only one child had an AQ-10 score of six,[2] and 90.3% had a score of three or lower. For children with ASD, 73.7% had an AQ-10 score of six or higher. A Mann-Whitney U test indicated that AQ-10 score was higher for children with ASD than for children in the control condition. Distribution of AQ-10 scores across groups is presented in Table 1.

Emotion Recognition Difficulties

Based on similar scales that were used as reference (e.g. Emotion Reactivity Scale; Nock, Wedig, Hilmberg, & Hooley, 2008), ten novel items assessed the emotion recognition difficulties of the children in themselves (e.g. "*describing his/her feelings is easy for my child*") and in others (e.g. "*my child has difficulties in recognizing emotions in others*"). Questions addressed the adequate recognition and interpretation of emotions, body language, facial expression, and tone of other people. Two items asked about the recognition and description of own emotions. The ten items are rated on a 4-point scale from 0 (*I do not agree at all*) to 3 (*I completely agree*) and had excellent internal consistency with a Cronbach's α = 0.95. The observed scale mean ranged from 0 to 3. Scores across groups are presented in Table 1.

Facial Recognition Task Material

Facial Stimuli

Facial stimuli were taken from the *Pictures of Facial Affect* (POFA[3]; Ekman 1976) database. The database includes 110 black and white pictures of 14 adult people who represent facial expressions of the six basic emotions: anger, disgust, fear, joy, sadness, and surprise. Additionally, a neutral face is included for every person. For the present study, three pictures were chosen from a licensed copy for the emotions joy, sadness, and anger respectively, resulting in a total of nine facial emotion expression items. Four pictures from two different males and five pictures of three different females were chosen as stimuli.

Music Stimuli

For the music stimuli, three music pieces were chosen in a pilot study with $N = 58$ students who listened to different music pieces to which they assigned an emotion. For each emotion, the music piece with the highest rating for the assigned emotion was chosen. Music pieces were rated on a 4-point scale on their emotional valence, with the joyful music piece receiving the highest rating on joy ($M = 3.62$; $SD = 0.06$), the sad music piece having the highest rating on sadness ($M = 3.10$; $SD = 0.76$), and the aggressive music piece resulting in the highest rating on anger ($M = 2.48$; $SD = 1.05$). Music pieces were provided by Altenmüller, Schürmann, Lim, and Parlitz (2002). The song chosen for the emotion joy was a Jazz-piece with rapid tempo and tones adequate to symbolize joyfulness (Dalla Bella, Peretz, Rousseau, & Gosselin, 2001). For sadness, a classical music piece played by a string quartet was chosen. This piece consisted of low tones and a slow rhythm representing sadness (Dalla Bella et al., 2001). Lastly, for anger, a music piece of the category psychedelic rock or heavy metal was used (Altenmüller et al., 2002). The stereo speakers of a laptop computer (see below) were used to play the songs at a moderate volume level (65 dB).

Facial Recognition Task

The facial recognition task consisted of nine trials, with one trial for every emotion depicted (anger, sadness, joy) and one trial for every music condition (neutral, congruent, or incongruent music). Before the actual emotion recognition task, every participant had a practice run to learn the instructions and understand the task. For the practice run, facial expression stimuli were replaced with a smiley depicting an emotion. The facial expressions were depicted on a black computer screen and the participants had to press one of three keys that were assigned to each of the three emotions, sadness, anger, and joy respectively. Participants were told to press the key as soon as they had identified the emotion. Collected data included the identity of the pressed key as well as reaction times (in milliseconds). Reaction times were assessed from the onset of the facial stimulus depiction to when a key was pressed. Facial stimuli were presented in 482 × 312 pixel format on a 17 inch laptop computer screen (HP Notebook – 17-×070ng – ENERGY STAR) and the songs were played at a moderate volume level (65 dB) through the internal speakers of the laptop.

Results

Data analyses were performed using IBM SPSS version 25. The criterion for statistical significance was set at $p < 0.05$. First, bivariate Spearman correlations were

[2] Excluding this participant from analyses did not change any results.

[3] Available at: https://www.paulekman.com/product/pictures-of-facial-affect-pofa/.

calculated for AQ-10 Child score, mean score for the scale for emotion recognition difficulty, sum of correctly recognized emotions through all trials (emotion recognition accuracy), and reaction times in the emotion recognition task. There was a significant correlation between AQ-10 scores and reaction times ($\rho = 0.31$, $p = 0.029$), indicating that higher autism spectrum quotients were associated with greater reaction times in the emotion recognition task. Also, there was a significant negative correlation between AQ-10 scores and the amount of correctly recognized emotions ($\rho = -0.31$, $p = 0.030$), meaning that higher autism spectrum quotients were related to fewer correctly recognized emotions in the facial recognition task. In addition, there was a strong and significant positive correlation between AQ-10 scores and the score on the parent-reported emotion recognition difficulties ($\rho = 0.86$, $p < 0.001$). Furthermore, there was a significant Pearson correlation between emotion recognition difficulties and reaction times in the emotion recognition task ($r = 0.31$, $p = 0.027$) and a significant negative correlation between emotion recognition scores and emotion recognition accuracy in the task ($r = -0.31$, $p = 0.027$). In a regression analysis, AQ-10 was also a significant negative predictor of emotion recognition accuracy, $b = -0.17$, $t(48) = 2.59$, $p = 0.013$, and a significant positive predictor of reaction times, $b = 195.23$, $t(48) = 3.32$, $p = 0.002$.

A one-way ANOVA was calculated with group as between-subjects factor (control children vs. children with ASD) and the emotion recognition accuracy as dependent variable. Levene's test of homogeneity of variances was significant ($p = 0.001$) and therefore Welch's test was used, $F(1, 22.54) = 14.01$, $p = 0.001$, $d = 1.29$. As expected, control participants had significantly better emotion recognition ($M = 8.48$, $SD = 0.81$) than participants with ASD ($M = 6.84$; $SD = 1.80$). Furthermore, a one-way ANOVA was calculated with group as between-subjects factor and total reaction times as dependent variable. Again, Levene's test of homogeneity of variances was significant ($p < 0.001$) and therefore Welch's Test was used, $F(1, 19.49) = 8.85$, $p = 0.008$, $d = 0.86$. As expected, control participants were significantly faster in recognizing emotions ($M = 1717.13$; $SD = 511.61$) than participants with ASD ($M = 3094.61$; $SD = 1978.44$). Lastly, a one-way ANOVA revealed a significant difference between groups on the scale for emotion recognition difficulties, $F(1, 48) = 49.72$, $p < 0.001$, $d = 2.11$. Again, children with ASD had significantly higher parent-reported emotion recognition difficulties than children in the control group. Mean scores for emotion recognition difficulties for both groups are provided in Table 1.

Additionally, a mixed measures ANOVA with Group (control vs. children with ASD) as between-subjects factor and reaction times (correct answers vs. incorrect answers) as within-subjects factor was calculated to see if reaction times differed between correct recognition and false recognition. The main within-subjects effect for reaction times was significant, $F(1, 23) = 37.44$, $p < 0.001$, $\eta^2 p = 0.62$, with participants taking significantly more time for correct recognition ($M = 2101.34$; $SD = 1349.12$) than for false recognition ($M = 655.78$; $SD = 660.04$). The interaction effect was not significant, $F(1, 23) = 0.85$, $p = 0.366$, $\eta^2 p = 0.04$. In this model with separated reaction times for correct and incorrect answers in the emotion recognition task, the between subjects effect was still significant, $F(1, 23) = 7.45$, $p = 0.012$, $\eta^2 p = 25$. This indicates that children with ASD need significantly more time in the emotion recognition task, independent of correct or incorrect recognition. However, neither total reaction time ($r = -0.23$, $p = 0.350$), nor reaction time for correct answers ($r = 0.06$, $p = 0.816$) were correlated to emotion recognition accuracy for children with ASD.

In a next step, a mixed measures ANOVA was calculated with Group as between-subjects factor (control participants vs. participants with ASD) and Music as within-subjects factor (no music vs. congruent music vs. incongruent music). Mauchly's test of sphericity was significant ($p < 0.001$) and therefore a method for the correction of degrees of freedom was applied. With the Greenhouse-Geisser correction ($\varepsilon = 0.71$) there was a significant within-subjects effect for Music, $F(1.42, 68.21) = 11.08$, $p < 0.001$, $\eta^2 p = 0.19$. Within-subjects contrasts revealed that there was no significant difference in the number of correctly recognized facial emotion expressions (emotion recognition accuracy) with no music ($M = 2.72$; $SD = 0.54$) or with congruent music ($M = 0.2.80$; $SD = 0.45$), $F(1,48) = 2.38$, $p = 0.129$, $\eta^2 p = 0.05$. However, there was a significant difference in emotion recognition between no music and incongruent music, $F(1, 48) = 9.28$, $p = 0.004$, $\eta^2 p = 0.16$, with emotion recognition being significantly better with no music ($M = 2.72$; $SD = 0.54$) than with incongruent music ($M = 2.34$; $SD = 1.02$). Lastly, participants with ASD were significantly better in recognizing facial emotion expressions (emotion recognition accuracy) with congruent music than with incongruent music, $F(1, 48) = 15.40$, $p < 0.001$, $\eta^2 p = 0.24$, see Fig. 1. More importantly, there was a significant interaction effect between Group and Music, $F(1.42, 68.21) = 5.11$, $p = 0.016$, $\eta^2 p = 0.10$. Separate contrasts for control participants and participants with ASD were calculated. There was only a significant difference in emotion recognition between congruent music and incongruent music for participants with ASD, $F(1, 18) = 9.32$, $p = 0.007$, $\eta^2 p = 0.34$, with participants with ASD having better facial emotion recognition accuracy with congruent music than with incongruent music. Also, participants with ASD had significantly better facial emotion recognition accuracy with no music than with

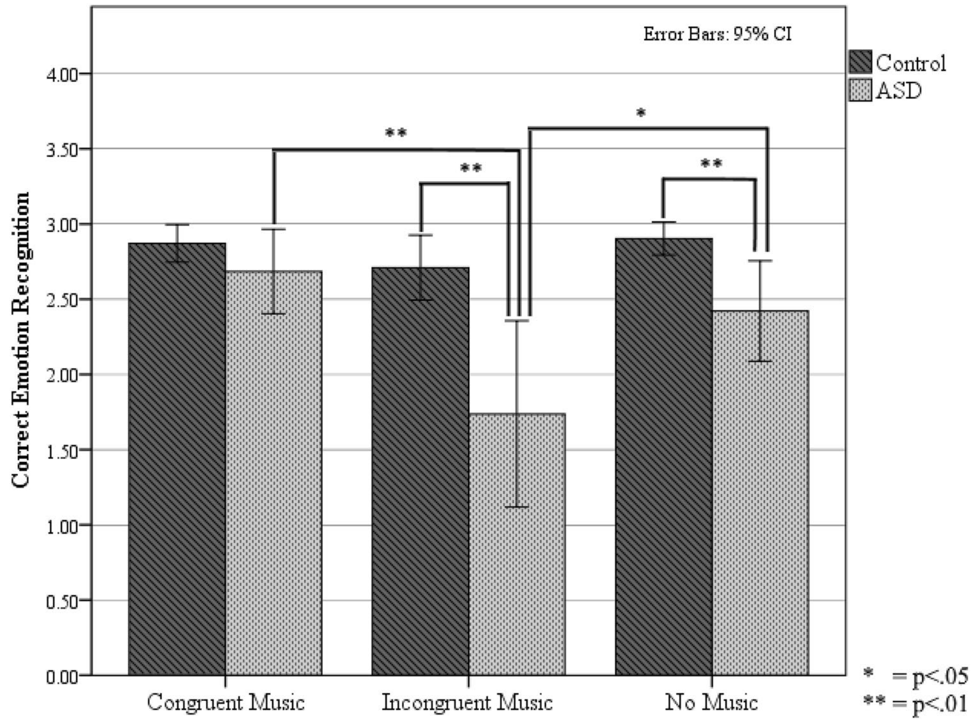

Fig. 1 Amount of correctly recognized facial emotion expressions for children with ASD and control children across music conditions

incongruent music, Welch's $F(1, 18) = 4.69$, $p = 0.044$, $\eta^2p = 0.21$, see Fig. 1. No other contrasts were significant.

There was also a significant group difference for emotion recognition with incongruent music, Welch's $F(1, 22.71) = 9{,}66$, $p = 0.005$, $d = 1.06$, with control participants having significantly better emotion recognition ($M = 2.71$; $SD = 0.59$) than participants with ASD ($M = 1.74$; $SD = 1.28$). Furthermore, there was a significant group difference for emotion recognition with no music, Welch's $F(1, 8.26) = 8.26$, $p = 0.009$, $d = 0.70$, indicating that control participants had significantly better emotion recognition with no music ($M = 2.90$; $SD = 0.69$) than participants with ASD ($M = 2.42$, $SD = 0.69$). Interestingly, there was no significant difference between groups for congruent music ($p = 0.215$).

To test if reaction times differed between children with ASD or control children in the correct emotion recognition across the different music conditions, a mixed measures ANOVA was calculated with group (control vs. children with ASD) as between-subjects factor and music (no music vs. congruent music vs. incongruent music) as within-subjects factor. As Mauchly's test of sphericity was significant ($p < 0.001$), the Greenhouse-Geisser correction of degrees of freedom was used. There was no significant main effect for music, $F(1.62, 69.71) = 1.48$, $p = 0.234$, $\eta^2p = 0.03$ and there was no significant interaction effect between group and music condition, $F(1.62, 69.71) = 0.50$, $p = 0.570$, $\eta^2p = 0.01$. This was confirmed in a within-subjects ANOVA only including children with ASD, $F(1.61, 20.89) = 0.38$, $p = 0.645$, $\eta^2p = 0.03$. Therefore, reaction times for the correct emotion recognition in faces did not differ across music conditions.

To test if the autism spectrum quotient is associated with diminished emotion recognition, bivariate Spearman correlations were computed with AQ-10 score and facial emotion recognition accuracy and reaction times for trials with congruent, incongruent, and no music. As expected, AQ-10 scores did not correlate with emotion recognition accuracy ($\rho = 0.06$, $p = 0.692$) in the congruent music trials. In contrast, AQ-10 scores were significantly and negatively associated with correctly recognizing emotional facial expressions in the incongruent music trials ($\rho = -0.29$, $p = 0.043$). In the trials with no music, the correlation between AQ-10 scores and emotion recognition was not significant ($\rho = -0.25$, $p = 0.078$). As expected, AQ-10 scores were also positively correlated with reaction times in the trials with no music ($\rho = 0.40$, $p = 0.004$), indicating that higher autism spectrum quotients were associated with greater reaction times in the emotion recognition task. Unexpectedly, however, the correlation between AQ-10 and reaction times of emotion recognition in the trials with congruent music was also significant ($\rho = 0.30$, $p = 0.037$) but not in the trials with incongruent music ($\rho = 0.08$, $p = 0.582$).

Further exploratory analyses were conducted. Firstly, a mixed-measures ANOVA was calculated with Group as between-subjects factor (control participants vs. participants with ASD) and the different depicted emotions in the emotion recognition task as within-subjects factor (anger vs. joy vs.

sadness). The analysis was computed to test for differences in emotion recognition difficulties between the three different emotions for participants with ASD and for control participants. There was no significant within-subjects effect for emotion, $F(2, 96) = 1.36$, $p = 0.262$, $\eta^2p = 0.03$, indicating that there was no significant difference in emotion recognition for angry, happy, or sad faces. Furthermore, the interaction between Group and emotion was also not significant, $F(2, 96) = 1.10$, $p = 0.338$, $\eta^2p = 0.02$.

Secondly, it was tested if emotion recognition varied across different ASD types. The ANOVA with emotion recognition as dependent variable and Group (control participants vs. early infantile autism vs. Asperger's vs. atypical ASD) as between-subjects factor was reanalyzed. Pairwise comparisons revealed that for emotion recognition there was only a significant difference between control participants and participants with early infantile autism, *Mdiff* = 2.28, *SE* = 0.59, $p = 0.002$, 95% CI [0.65, 3.92], and between participants with atypical ASD, *Mdiff* = 2.15, *SE* = 0.55, $p = 0.002$, 95% CI [0.64, 3.67]. The difference between control participants and participants with Asperger's Syndrome was not significant, *Mdiff* = 0.85, *SE* = 0.49, $p = 0.512$, 95% CI [−0.49, 2.21]. Adjustments for multiple comparisons were done with the Bonferroni correction. However, these results must be interpreted with caution as group sizes were very small and therefore analyses could have lacked statistical power. Only $n = 5$ participants had early infantile autism, $n = 8$ participants had Asperger's Syndrome and $n = 6$ participants had atypical ASD.

Additionally, age was not a significant predictor of emotion recognition difficulties, neither for the whole sample ($N = 50$), $b = -0.09$, $t(48) = -0.88$, $p = 0.384$, nor for children with ASD ($n = 19$), $b = -0.06$, $t(17) = -0.61$, $p = 0.553$. Age was also not a significant predictor of emotion recognition accuracy for children with ASD, $b = 0.36$, $t(17) = 1.31$, $p = 0.207$. For the whole sample, age was only a marginally significant predictor of emotion recognition accuracy, $b=0.29$, $t(48) = 1.82$, $p = 0.075$. Furthermore, gender was neither a significant predictor of emotion recognition accuracy for the whole sample ($N = 50$), $b = -0.31$, $t(48) = -0.71$, $p = 0.482$, nor for children with ASD ($n = 19$), $b = -0.25$, $t(17) = -0.28$, $p = 0.780$.

Discussion

The present study tested the influence of emotion congruent, incongruent, or no music on facial emotion recognition between children without ASD and children with ASD. In accordance with the literature, children without ASD were hypothesized to have greater facial emotion recognition accuracy than children with a diagnosed ASD. Furthermore, it was hypothesized that for congruent music, facial emotion recognition would be improved for children with ASD compared to incongruent or no music played in the background. As there is no difference in emotion recognition in music between children with ASD and neurotypical children (Caria et al., 2011), it was expected that with congruent emotional music, children with ASD would have similar emotion recognition accuracy as the children in the control group. The present findings both confirm and extend the existing literature.

As expected, children without ASD generally had significantly greater facial emotion recognition accuracy in the non-congruent and the no music conditions than children with ASD in the non-congruent and the no music condition. The control children also needed significantly less time for emotion recognition than children with ASD did in the mentioned conditions. However, when congruent emotional music was played in the background, the difference in the recognition accuracy of the corresponding facial emotion between children in the control group and children with ASD was no longer present. Emotion recognition accuracy of children without ASD was not influenced by music. In contrast, even though there was no significant difference for congruent music for children with ASD, compared to the control children, children with ASD had significantly more difficulties in emotion recognition when faces appeared with incongruent music or no music at all. This suggests a facilitating effect of congruent music on facial emotion recognition ability for children with ASD. These results point to a distinct pattern of cross modal cueing among children with ASD and control children, demonstrating the advantage of simultaneous congruent music in the recognition of facial emotional expressions for children with ASD. Children with ASD may be able to extract the type of patterns of psychoacoustic cues that convey emotions in music described by Juslin (2013), which explains how music and emotions are interconnected. Furthermore, this may provide an explanation to the reason why the accuracy of ASD participants dropped in the incongruent condition compared to the condition without music, while such effect did not occur for the control children. Children with ASD may, in fact, be more strongly influenced by the auditory emotional cues provided by music than neurotypical children and rely more strongly on that information.

Children with ASD needed significantly more time for emotion recognition. However, even when they made a mistake, children with ASD needed more time. This is in line with findings that children with ASD use more deliberate strategies to recognize emotion expressions (Rump et al. 2009; Tracy et al. 2011). Music did not improve recognition speed for children with ASD, and the autism spectrum quotient was positively correlated with recognition time when congruent or no music was played and was not correlated with reaction time when incongruent music was played. For children with high autism spectrum quotients, congruent music and no music may prompt them to focus more on emotion recognition strategies, resulting in greater reaction times. Studies have shown that when individuals with ASD take more time, they are as accurate as neurotypical individuals in emotion recognition (e.g. Piggot et al., 2004). In the present study, however, reaction times were not correlated with emotion recognition accuracy in children with ASD. Children with ASD did not use more time to be more accurate in the emotion recognition task. Looking at the full picture, results seem to support the findings that individuals with ASD generally need more time for emotion recognition (Bal et al., 2010; Capps, Yirmiya, & Sigman, 1992; Homer & Rutherford, 2008).

Facial emotion recognition deficits are a fundamental characteristic of ASD. The present results of a diminished facial emotion recognition accuracy in children with ASD confirm this. We also found that higher AQ-10 scores predicted lower emotion recognition accuracy and greater reaction times during the task. They were also strongly associated with emotion recognition difficulties reported by parents. As the AQ-10 child is highly predictive of ASD (Allison et al., 2012) and has high discriminative validity for ASD (Booth et al., 2013), the present results are in line with previous findings on the difficulties of children with ASD to recognize emotions. However, it must be noted that only 73.7% of participants with ASD had an AQ-10 score above the threshold of 6. Five of the 19 children diagnosed with ASD had a value below 6, with three of these five children even having a score below 5. This could be due to the fact that the present sample consisted of children diagnosed with different subtypes of ASD according to DSM-IV criteria (American Psychiatric Association, 2000).

The present finding of a comparably worse emotion recognition performance in children with ASD is in line with the existing literature. People with ASD have significantly worse emotion recognition of the basic emotions *anger*, *disgust*, *fear*, *sadness*, *surprise,* and even to a less extent for *joy* than neurotypical individuals (Lozier et al. 2014; Uljarevic & Hamilton, 2013). This could be due to differences in brain structures in patients with ASD. For example, the amygdala, which is in part responsible for emotion processes like recognition, is impaired in people with ASD (Nacewicz et al., 2006) and is important for the modulation of social behavior (Amaral, 2003). Individuals with ASD have a deviant facial emotion perception, as they do not tend to focus on the eye area of the face (Nacewicz et al., 2006). This has an impact on their social interactions and communication. Unlike facial emotion perception, musical cues are processed differently and therefore may not be impaired in ASD. Even more so, there is evidence that the use of music can enhance facial emotion recognition accuracy (Brown, 2017; Katagiri 2009; Logeswaran & Bhattacharya, 2009).

Unlike other studies (e.g. Tracy et al., 2011; Uljarevic & Hamilton, 2013), the present study did not find any differences in emotion recognition for the three emotions, joy, sadness, and anger. For example, Uljarevic and Hamilton (2013) provided evidence that joy is perceived significantly better in facial expression than other emotions and that there is no difference between neurotypical and individuals with ASD for the recognition of happiness. This was not the case in the present study, which only comprised 9 trials with just 3 pictures showing happy faces. Also, unlike other studies (e.g., Lozier et al., 2014), the present study did not find any relationship between age and emotion recognition difficulties or emotion recognition accuracy for children with ASD. However, it must be noted that their age did not vary much in the present sample, as children were only between 8 and 12 years old. Furthermore, the autism spectrum quotient was used to cross-validate the group results. Additionally, the pattern of performance found in the current study is not in line with research that has found decreased arousal to music in children with ASD (Stephenson et al., 2016a) and thus does not support the claim that physiological effects of music-evoked emotions may be lower in childhood than in adulthood in ASD (Stephenson et al., 2016b).

The present study substantially extends the existing research. Firstly, unlike Brown (2017), for example, we did not focus on gradual differences in the performance of children with ASD with regard to perceiving sad faces as sadder or happy faces as happier. Instead, facial emotion recognition accuracy was assessed directly, that is, qualitatively. This represents an important and socially relevant extension of the literature as this allows for testing the direct influence of musical stimuli on facial emotion recognition accuracy. For example, it allows to test if interventions that use congruent music in emotion recognition trainings are effective in improving accuracy. Improving emotion recognition accuracy for facial expressions will have benefits for children with ASD in their social interaction and communicative behavior. Furthermore, in the present study not only children with high-functioning ASD were included but children with different types of ASD which allows for a more fine-grained interpretation of results. Based on the present findings, we may speculate that only participants with early or atypical ASD, but not those with Asperger's Syndrome, may benefit from the performance boosting effect of music that is congruent with the facial emotion.

Some limitations to the present study must be noted. Firstly, the facial recognition task was constructed using only one trial for each category (i.e. the three emotions each with congruent vs. incongruent vs. no music). Although this decision was consciously made against the background of not cognitively overburdening children diagnosed with ASD, this could have limited the power of statistical analyses as the opportunity for variance differences was smaller than when more trials would have been used. However, in fact, in a similar study by Brown (2017), it was argued against using an additional control condition with no music due to concerns of overloading children with ASD with too many stimuli. As the control condition without music was essential for interpretation of results, the number of stimuli was limited instead.

Additionally, as the current study was designed as a field study, no true systematic control of environmental confounds was possible. Furthermore, the sample size was not large enough to ensure valid results for the comparison of children with different ASD types. Even though the current version of the DSM-5 (American Psychiatric Association, 2013) does not distinguish anymore between the different types of autism but understands autism as a developmental disorder that lies on a spectrum (i.e., the autism spectrum disorder), the differentiation between different forms of autism could be useful to determine the specificity of effects. However, using DSM-IV criteria that resulted in having many "mild" forms of ASD might have limited the statistical power. A different option in the future would be to use DSM-5 criteria for ASD and having a measure of the extent of ASD for participants (e.g., using measures like the AQ-10). However, even when considering the lowered statistical power, it could be interesting to follow up on the results on the different subtypes of ASD in future studies.

In conclusion, the present findings show that music affects how children with ASD process emotional faces. This has potential implications for intervention designs. Music interventions have already shown promising effectiveness in increasing social skills in people with ASD (Geretsegger et al., 2014). If playing congruent emotional music to enhance emotion recognition skills proves effective, it could be integrated into interventions like music therapy for children with ASD. When emotion recognition does not differ between neurotypicals and individuals with ASD when congruent music is played, then it would be promising to use music to increase the salience of emotions for people with ASD and to help them improve their emotion recognition skills. As a consequence, social interactions may become easier for people with ASD. Future studies are needed to assess if music can have a long-lasting effect on emotion processing.

Acknowledgments The authors have no external funding sources or conflicts of interest to declare.

Author Contributions Gary L. Wagener, Andreia Costa, and André Melzer are the main authors of manuscript. Madeleine Berning, Andreia Costa, and André Melzer contributed in study design and execution. Gary L. Wagener contributed in data handling and analysis. André Melzer and Georges Steffgen contributed in supervision.

Compliance with Ethical Standards

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

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

Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief "red flags" for autism screening: The short autism spectrum quotient and the short quantitative checklist in 1000 cases and 3000 controls. *Journal of the American Academy of Child & Adolescent Psychiatry, 51*(2), 202–212.e7. https://doi.org/10.1016/j.jaac.2011.11.003.

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 lateralization patterns. *Neuropsychologia, 40*, 2242–2256. https://doi.org/10.1016/s0028-3932(02)00107-0.

Amaral, D. G. (2003). The amygdala, social behavior, and danger detection. *Annals of the New York Academy of Sciences, 1000*, 337–347. https://doi.org/10.1196/annals.1280.015.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., Text Revision). Washington, DC: Author.

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

Bal, E., Harden, E., Lamb, D., Van Hecke, A. V., Denver, J. W., & Porges, S. W. (2010). Emotion recognition in children with autism spectrum disorders: Relations to eye gaze and autonomic state. *Journal of Autism and Developmental Disorders, 40*(3), 358–370. https://doi.org/10.1007/s10803-009-0884-3.

Baron-Cohen, S. (1995). *Learning, development, and conceptual change. Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: The MIT Press.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*, 5–17. https://doi.org/10.1023/a:1005653411471.

Booth, T., Murray, A. L., McKenzie, K., Kuenssberg, R., O'Donnell, M., & Burnett, H. (2013). An evaluation of the AQ-10 as a brief screening instrument for ASD in adults. *Journal of Autism and Developmental Disorders, 43*, 2991–3000. https://doi.org/10.1007/s10803-013-1844-5.

Brown, L. S. (2017). The influence of music on facial emotion recognition in children with Autism Spectrum Disorder and neurotypical children. *Journal of Music Therapy, 54*(1), 55–79. https://doi.org/10.1093/jmt/thw017.

Capps, L., Yirmiya, N., & Sigman, M. (1992). Understanding of simple and complex emotions in non-retarded children with autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines,* *33*, 1169–1182. https://doi.org/10.1111/j.1469-7610.1992.tb00936.x.

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*, 2838–2849. https://doi.org/10.1093/cercor/bhr084.

Dalla Bella, S., Peretz, I., Rousseau, L., & Gosselin, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition, 80*(3), B1–B10. https://doi.org/10.1016/s0010-0277(00)00136-0.

Fink, E., de Rosnay, M., Wierda, M., Koot, H. M., & Begeer, S. (2014). Brief report: Accuracy and response time for the recognition of facial emotions in a large sample of children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 44*(9), 2363–2368. https://doi.org/10.1007/s10803-014-2084-z.

Gebauer, L., Skewes, J., Westphael, G., Heaton, P., & Vuust, P. (2014). Intact brain processing of musical emotions in autism spectrum disorder, but more cognitive load and arousal in happy vs. sad music. *Frontiers in Neuroscience, 8*, 192. https://doi.org/10.3389/fnins.2014.00192.

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

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

Heaton, P. (2009). Assessing musical skills in autistic children who are not savants. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1522), 1443–1447. https://doi.org/10.1098/rstb.2008.0327.

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. https://doi.org/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*, 1405–1410. https://doi.org/10.1017/s0033291799001221.

Homer, M., & Rutherford, M. D. (2008). Individuals with autism can categorize facial expressions. *Child Neuropsychology, 14*, 419–437. https://doi.org/10.1080/09297040802291715.

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

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. https://doi.org/10.1093/jmt/46.1.15.

Kuusikko, S., Haapsamo, H., Jansson-Verkasalo, E., Hurtig, T., Mattila, M. L., Ebeling, H., Jussila, K., Bölte, S., & Moilanen, I. (2009). Emotion recognition in children and dolescents with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*, 938–945. https://doi.org/10.1007/s10803-009-0700-0.

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

Lozier, L. M., Vanmeter, J. W., & Marsh, A. A. (2014). Impairments in facial affect recognition associated with autism spectrum disorders: A meta-analysis. *Development and Psychopathology, 4*(1), 933–945. https://doi.org/10.1017/S0954579414000479.

Molnar-Szakacs, I., & Heaton, P. (2012). Music: A unique window into the world of autism. *Annals of the New York Academy of Sciences, 1252*(1), 318–324. https://doi.org/10.1111/j.1749-6632.2012.06465.x.

Nacewicz, B. M., Dalton, K. M., Johnstone, T., Long, M. T., & McAuliff, E. M. (2006). Amygdala volume and nonverbal social impairment in adolescent and adult males with autism. *Archives of General Psychiatry, 62*(12), 1417–1428. https://doi.org/10.1001/archpsyc.63.12.1417.

Nock, M. K., Wedig, M. M., Hilmberg, E. B., & Hooley, J. M. (2008). The emotion reactivity scale: Development, evaluation, and relation to self-injurious thoughts and behaviors. *Behavior Therapy, 39*, 107–116. https://doi.org/10.1016/j.beth.2007.05.005.

Piggot, J., Kwon, H., Mobbs, D., Blasey, C., Lotspeich, L., Menon, V., Bookheimer, S., & Reiss, A. L. (2004). Emotional attribution in high-functioning individuals with autistic spectrum disorder: A functional imaging study. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*(4), 473–480. https://doi.org/10.1097/00004583-200404000-00014.

Rump, K. M., Giovannelli, J. L., Minshew, N. J., & Strauss, M. S. (2009). The development of emotion recognition in individuals with autism. *Child Development, 80*(5), 1434–1447. https://doi.org/10.1111/j.1467-8624.2009.01343.x.

Shi, Z.-M., Lin, G.-H., & Xie, Q. (2016). Effects of music therapy on mood, language, behavior, and social skills in children with autism: A meta-analysis. *Chinese Nursing Research, 3*, 137–141. https://doi.org/10.1016/j.cnre.2016.06.018.

Stephenson, K. G., Quintin, E. M., & South, M. (2016a). Age-related differences in response to music-evoked emotion among children and adolescents with autism spectrum disorders. *Journal of autism and developmental disorders, 46*(4), 1142–1151. https://doi.org/10.1007/s10803-015-2676-2.

Stephenson, K. G., Quintin, E. M., & South, M. (2016b). Common threads, age-related differences, and avenues for future research: Response to Heaton. *Journal of Autism and Developmental Disorders, 46*(4), 1494–1497. https://doi.org/10.1007/s10803-016-2750-4.

Tracy, J. L., Robins, R. W., Schriber, R. A., & Solomon, M. (2011). Is emotion recognition impaired in individuals with autism spectrum disorder? *Journal of Autism and Developmental Disorders, 41*, 102–109. https://doi.org/10.1007/s10803-010-1030-y.

Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: A formal meta-analysis. *Journal of Autism and Development Disorders, 43*(7), 1517–1526. https://doi.org/10.1007/s10803-012-1695-5.