Learning and Instruction 45 (2016) 85–92

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

Effects of music and natural science training on aggressive behavior

Ingo Roden [a, c, *], Florian D. Zepf [e, f], Gunter Kreutz [b, c], Dietmar Grube [a], Stephan Bongard [d]

[a] Educational Psychology, Department of Educational Sciences, Carl von Ossietzky University Oldenburg, 26129 Oldenburg, Germany
[b] Department of Music, School of Linguistics and Cultural Studies, Carl von Ossietzky University Oldenburg, 26129 Oldenburg, Germany
[c] Speech and Music Lab, Carl von Ossietzky University Oldenburg, 26129 Oldenburg, Germany
[d] Department of Psychology, Goethe-University Frankfurt, 60323 Frankfurt am Main, Germany
[e] Centre & Discipline of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, School of Psychiatry and Clinical Neurosciences & School of Paediatrics and Child Health, Faculty of Medicine, Dentistry and Health Sciences, The University of Western Australia, Perth, WA, Australia
[f] Specialised Child and Adolescent Mental Health Services (CAMHS), Department of Health in Western Australia, Perth, WA, Australia

ARTICLE INFO

Article history:
Received 5 August 2015
Received in revised form
27 May 2016
Accepted 8 July 2016

Keywords:
Music training
Natural science training
Aggressive behavior
Cardiovascular and neurohumoral stress response
Transfer

ABSTRACT

Extended music lessons have been suggested to reduce stress responses, and to increase well-being in primary school children. We investigated this assumption with regard to the provocation of aggressive behavior in primary school children ($N = 34$; 7–8 years of age). A computerized modified version of the Point-Subtraction Aggression Game ('Stimulated Aggression by Virtual Opponent'; SAVO) was used in this sample. Self-report (Positive and Negative Affect Schedule, PANAS) and physiological measures including systolic blood pressure (SBP), diastolic blood pressure (DBP), heart rate (HR), and saliva cortisol concentrations were recorded before, during, and after the SAVO task. For the following 18 months, one group of children received weekly sessions of extended instrumental music lessons ($n = 14$; music group), while a control group received natural science training ($n = 20$; control group). A set of repeated measures analyses of variance (ANOVAs) did not show any differences in physiological measures between groups. Moreover, only children in the control group, but not music children, showed a significant increase of reactive aggressive behavior after the SAVO task. These results suggest that music training positively modulates reactive aggressive behavior in primary school children.

1. Introduction

Learning to play a musical instrument during childhood has been associated with a wide range of academic, cognitive, emotional, personal and social benefits (Hallam, 2010). Such reported benefits include general measures of intelligence (Schellenberg, 2004), attention and processing speed (Roden, Könen, et al., 2014), working memory (Lee, Lu, & Ko, 2007; Roden, Grube, Bongard, & Kreutz, 2014), academic achievements (Anvari, Trainor, Woodside, & Levy, 2002; Southgate & Roscigno, 2009), and aspects of auditory memory (Ho, Cheung, & Chan, 2003; Roden, Kreutz, & Bongard, 2012). Recent studies also suggest that transfer effects might well extend to individual and socials skills, such as self-esteem (Costa-Giomi, 2004; Rickard et al., 2013), migrants' acculturation processes (Frankenberg et al., 2014), and emotional sensitivity (Thompson, Schellenberg, & Husain, 2004). The present study was designed to investigate music-induced transfer effects on children's reactive aggressive behavior and psychophysiological stress responses in primary school.

In a more general manner, stress is defined as 'an emotional state caused by various emotional or circumstantial factors' (Yehuda, 2011, p. 86). Importantly, prolonged states of stress may show a multitude of physical and behavioral manifestations. Many of these can be harmful to the person itself, as well as to others. Therefore, coping and stress management has become a prevailing topic in both theoretic and applied psychological research.

1.1. Music listening/training and aggressive behavior

Musical behaviors, in general, and music listening, in particular, have proven to mediate (calming or intensifying) individuals' stress responses (Linnemann, Ditzen, Strahler, Doerr, & Nater, 2015; Sandstrom & Russo, 2010). For example, music listening has been shown as an effective means to reduce agitated behavior and aggression in patients suffering from dementia (e.g., Koger, Chapin,

* Corresponding author. Educational Psychology, Department of Educational Sciences, Carl von Ossietzky University Oldenburg, 26129 Oldenburg, Germany.
E-mail address: Ingo.roden@uni-oldenburg.de (I. Roden).

http://dx.doi.org/10.1016/j.learninstruc.2016.07.002
0959-4752/

& Brotons, 1999) as well as in school-children with learning difficulties (Savan, 1999). In the latter study, decreases of attention-seeking behaviors as well as cardiac responses (heart rate, systolic and diastolic blood pressure) were observed when selected chamber music by Mozart was played as background during regular class sessions.

A systematic review by Varela, Abrami, and Upitis (2014) suggests a positive contribution of music instruction on self-regulation. Therefore, learning to play a musical instrument rather than mere listening to music also contributes to modulate psychophysiological stress, anxiety, as well as self-regulatory and reactive emotional behavior, particular in school-aged children. For example, Choi, Lee, and Lee (2010) found that primary-school children (10–12 years of age) showed reduced levels of aggression as well as enhanced levels of self-esteem after 15 weeks of a comprehensive music program including singing, making and playing musical instruments, and song-writing. Similarly, Ho, Tsao, Bloch, and Zeltzer (2011) found that low-income children benefitted from drumming groups across a range of socio-emotional problems. In addition, Costa-Giomi (2004) showed that formal piano lessons over a period of three years both enhanced global self-esteem and reduces aggressive behavior in primary school children. Laohawattanakun et al. (2011) showed that adolescents (15–17 years of age) receiving music training were less vulnerable to examination-induced stress as measured by salivary cortisol concentrations. Furthermore, Lindblad, Hogmark, and Theorell (2007) reported preliminary data on the effects of one additional hour per week of music education for 5th-6th graders over three time points within a period of one school year. Comparing the music group to controls, the authors found a significant decrease in afternoon cortisol levels only for the music training group at the third time point of assessment. Contrary to their hypothesis, no differences were detected in social anxiety (in the context of the children's peer relations) or in the description of behavioral or emotional problems as assessed in the parental version of the Child Behavior Check List (CBCL, Achenbach, 1991) across groups. Finally, Rickard et al. (2013) reported positive effects of music training on children's global and social self-esteem, that appeared as a protection against the age-related decline in self-esteem in the early years of primary school children (e.g., Burnett, 1996).

Despite some positive results, the overall pattern of socio-emotional effects of music training is inconclusive. For instance, Rickard, Bambrick, and Gill (2012) examined nearly 250 primary and secondary school children to investigate the effects of school-based music training on cognitive skills, self-esteem and reactive aggressive behavior via standardized questionnaires. In contrast to their expectations, no substantial benefits from music training were apparent. Furthermore, Knocks-Anderson and Rickard (2007) found no evidence for decreasing anger expression or superior self-esteem skills in musically trained school children either. Studies on the effects of music training on psychological and physiological measures of stress responses and aggressive behavior in primary school children are still rare (see Varela et al., 2014). Furthermore, because different approaches of measuring reactive aggressive behavior and stress responses in children were used as indexed by (subjective) self-rating scales or (more objective) tasks makes it difficult to compare the obtained results of the reported studies. Especially whether music training may modulate levels of reactive aggressive behaviors and stress in primary school children remains controversial. Furthermore, research reported here mostly focused on private music interventions rather than on school-based music lessons, and most of the reported studies only included control groups without any equivalent training compared to the intervention.

1.2. Aims and hypotheses

The aims of the present study were to investigate the impact of a school-based instrumental music training program on the provocation of reactive aggressive behavior in primary school children over a period of 18 months, and to compare the music children's reactions to children of the same age and social background, who received natural science training (here: control group). At two time points, measures of reactive aggressive behavior, cardiovascular (SBP, DBP and HR) and neurohumoral (cortisol) stress responses were obtained, as well as self-reported positive and negative affect ratings (PANAS), before, during, and after a provocation task. We hypothesized that children receiving instrumental music training respond to a provocation task with significantly less aggressive behavior than children receiving natural science training. Moreover, we investigated whether children in the music group showed lower psychological, cardiovascular and neurohumoral stress responses to the provocation of reactive aggressive behavior when compared to children in the natural science group.

2. Materials and methods

2.1. Schools and participants

A total of 34 children (mean age at the onset of the study = 7.76 years; *SD* = 0.74; 15 males, 19 females) participated in this study. According to a pre-study power-analysis conducted using G*Power (Erdfelder, Faul, & Buchner, 1996) this sample size was considered sufficient to proof small to medium effects ($f = 0.25$) in a mixed within/between subjects design with two time points (α: 0.05, power (1-β): 0.80, correlations between repeated measures: $r = 0.50$). Participants were recruited from six primary schools located in different parts of Germany. Fourteen children participated in an instrumental music training program (music group), and the remaining 20 children took part in a natural science training program and served as a control group. The following instruments were played by the children in the music group: 4 × guitar, 3 × violin, 1 × cello, 4 × flute, 1 × trumpet and 2 × keyboard). One child played two instruments (keyboard and violin). Gender was well balanced across groups ($\chi^2 = 0.68$; $p = 0.41$).

The musical training program was guided by a foundation called “Jedem Kind ein Instrument” [An instrument for every child] (see www.jedemkind.de for further details) and was established in over 650 primary schools in the German federal state North Rhine-Westphalia. The natural science training was part of nationwide large-scale programs of science education in Germany, including more than 850 primary schools (Ostermeier, Prenzel, & Duit, 2010).

Children had already been assigned to the music and the natural science training program prior to this study. Hence age, class, school grades for social behavior, migration background, parents' educational achievements, household income and cultural practice were assessed in order to control for potential baseline (T1) differences. Furthermore, we measured the amount of extra-school-activities for every participant, which was well balanced between groups ($\chi^2 = 0.79$; $df = 2$; $p = 0.37$). To minimize school or classroom effects, the children in the music group were randomly chosen out from six different classes from three different schools, whereas children in the natural science training program were chosen out from three different classes of three schools. The study was approved by the institutional review boards of the universities of Frankfurt am Main and Oldenburg in Germany. Additional written informed consent was obtained from school administration, parents and children as well.

2.2. Intervention

Children in the music group received extra weekly training of 45 min on musical instruments of their choice in addition to the regular school music curriculum. The training included pitch, rhythm and singing exercises at basic level. Lessons were organized in small groups with a maximum of five students at school. All children in the music group obtained professional music lessons by trained instrumental teachers from public music schools. Children in the control group received extended education in natural science that was embedded into the regular school curriculum as an extension in content, but not in teacher-pupil contact time. The format of this natural science extension consisted of, similar as in the music intervention, 45-min sessions of intensive work in small groups under the supervision of specially-trained teachers. Here, emphasis was placed on students' collaborative learning and social communication of mathematics and natural science skills. Children in both groups were tested individually in a separated room in the course of one hour at two time points (T1, T2) about 18 months apart. Data of the individual testing were ascertained between 8 a.m. and 12 p.m. at both time points. Children of the control group received only regular curriculum based music lessons at school.

2.3. Task

A computerized competitive reaction time game called 'Stimulated Aggression by Virtual Opponent' (SAVO[1]) was administered in this study to operationalize reactive aggressive behavior. The software is based on the "Point Subtraction Aggression Game" (PSAG) by Pelham et al. (1991), which is workable for minors. Participants were instructed to play the game against a fictively opponent of the same gender and age linked via a network scenario. Their task was to press a button as soon as possible whenever the image of a "soccer-ball" appeared on a computer screen stylized as a soccer field. When a participant was faster than his or her opponent he or she won the trial and gained a fixed amount of 50 points to their total score. Furthermore, the winner of a trial was allowed to decrease to their own choice the total score of their opponent by 0–100 points. Whenever a participant lost a trial, the opponent removed points from their total score. Indeed, there was no real opponent, and the sequences of losing and winning trials was predetermined and fixed.

For any trial lost by the participants (22 out of 48) the amount of deducted points varied between 20 and 40 points, representing a low provocation rate (see also Zepf et al., 2008). The amount of points participants subtracted from their fictional opponents account directly following a trial they had lost themselves served as measure of provoked reactive aggressive behavior. In addition, reaction time (RT) and decision time (DT) for the 26 winning trials were analysed as measure of the participants state off alertness (RT) and decidedness (DT). Reaction time was defined by the amount of time the participants needed to press the button after the "soccer ball" appeared on the screen. Decision time was defined as the time between the presentation of the information that the participant has won the trial and the time when the amount of point reductions was entered into the computer.

2.4. Questionnaires

We used standardized questionnaires for demographic and socioeconomic variables. Items were taken from a standardized questionnaire of the "Progress in International Reading Literacy Study [PIRLS]" (Bos et al., 2005). This information was acquired via questionnaires and telephone interviews with the parents as part of the baseline measurements of a larger study and assessed by a Rasch-scaled composite score of 14 variables including parental education, income and cultural practice (see Bos et al., 2005). Feedback about social behaviors was given via school marks by the teachers. Musical background was assessed with particular respect to extra-curricular activities, including instrumental music training in the control group. Children in the music group were musically naïve and did not play an instrument before the intervention started. Moreover, none of the children in the control group had played an instrument before or during the study.

Affective stress responses were assessed using an altered version of the German adaptation of the Positive and Negative Affect Schedule (PANAS; Krohne, Egloff, Kohlmann, & Tausch, 1996; Watson, Clark, & Tellegen, 1988). To strengthen the comprehensibility for primary school children, only 10 out of 20 adjectives were used in the current study, and the rating scale used three instead of five points. Adjectives selected to represent positive affect were *alert, pride, happy, attentive* and *excited*. Adjectives used for negative affect ratings were *angry, afraid, nervous, sad* and *guilty*. The intensity of each emotional experience was reported on a three-point-scale ranging from "not at all" to "a lot". Coefficient alpha for positive affects was 0.89 and 0.85 for negative affects. Retest reliabilities of positive affects were $r_{tt} = 0.54$ and $r_{tt} = 0.45$ for negative affects.

2.5. Physiological measures

Cardiovascular measures of heart rate (HR), systolic (SBP) and diastolic (DBP) blood pressure and a neurohumoral measure (salivary cortisol) were taken to determine the physiological stress response. SBP, DBP (mmHg) and HR (bpm) were measured using semi-automatic oscillometric blood pressure devices (boso-medicus prestige, Bosch & Sohn GmbH u. Co. KG, Jungingen, Germany). Salivary cortisol was collected with Sarstedt Salivettes® – a device that consists of a cotton wool swab within a plastic tube. Participants were instructed to open the plastic tube and to take the swap out into their mouth without touching it with their hands. After moving it through their mouth for about one minute participants were told to place the swab back into the plastic tube. Salivary samples were kept at −30 °C until assayed. Saliva analyses were conducted at the Department of Biological Psychology, Technical University of Dresden, Germany.

2.6. Procedure

The tests were conducted at the elementary schools attended by the participants of both experimental and control groups in the morning hours during class time. Children were tested individually in a separate room, which was acoustically shielded from the classrooms. Concerning the hormone sampling, it must be considered that in the morning, cortisol concentrations usually show a strong decrease. Since elementary schools in Germany teach only in the morning we were restricted to this time period. However, in order to exclude differences in time of the day as sources for possible differences the time when starting the protocol was registered and subsequent analyses indicated no significant difference between groups [$t(32) = 1.95$, $p > 0.05$] or between the time points T1 and T2 [$t(33) = 1.94$, $p > 0.05$] in time of day when the session started.

The SAVO task was administered, and cardiovascular (SBP, DBP and HR) and neurohumoral (salivary cortisol) measures where taken by trained research assistants.

[1] SAVO is a developed in-house software platform that allows for the setup of the PSAG.

Table 1
Overview of the design of experiments for the measured dependent variables at both time points T1 and T2.

Sequence	Phase	HR and BP taken at …	PANAS	Saliva
Beginn				
	Pre-Stress	1 min 3 min 5 min	I	I (pre-stress-value)
SAVO-Start				
	Task A	After the first award of points2 min after the first award of points		
	Break (3 min)	1.5 min	II	
	Task B	After the first award of points 2 min after the first award of points		
SAVO-End				
	Post-Stress A	1 min 3 min		II (stress-value)
Painting (10 min)				
	Post-Stress B	1 min 3 min after painting	III	III (post-stress-value)
End				

The order of the measures was fixed and equal at both time points and for each child. Breaks were determined for all participants and defined in the research design protocol (see Table 1). The standardized sociodemographic and personality related questionnaires were completed on a different day prior to this stress protocol.

The stress protocol of the study was arranged in three different phases (pre-stress, stress, and post-stress periods) at both time points (T1 and T2). Phase one (pre-Stress) covered three readings of HR, SBP and DBP, the collection of one saliva sample and the completion of the PANAS.

The pre-stress period was followed by the stress-period, which was split into two halves (Task A and Task B). Over this stress-period participants took part in the SAVO task and while they were playing that game HR, SBP and DBP were obtained after the first award of points was given and again two minutes thereafter. During a short break of three minutes in-between the two halves of the stress-period the PANAS was completed and saliva sample was obtained. Once, this break was over, the second half of the stress-period (Task B) started. Like before SBP, DBP and HR were measured after the first award of points had been given and two minutes later. After the SAVO task was completed phase three started (post-stress). SBP, DBP, HR and salivary cortisol levels were assessed twice and another PANAS was accomplished. After these first post-stress readings the children were asked to draw a picture for at least ten minutes of time. Thereafter, SBP, DBP and HR were taken again twice. Then the stress-protocol was finished. The children were disconnected from the equipment, thanked and escorted back to their classroom. The assessment from transferring each child to the experimental room and returning back to the classroom lasted approximately 45 min.

2.7. Data handling

Before analyses, data cleansing of the psychophysiological and the SAVO task measures were carried out. Ranges for acceptable physiological measures were defined as follows: HR (50–180 beats per minute), SBP (60–180 mmHg) and DBP (30–125 mmHg). Data beyond these limits were considered as artefacts and were coded as missing values. Concerning the neurohumoral cortisol measures no outliers were found. For each period of the protocol (pre-stress, stress, post-stress) the readings for SBP, DBP and HR were averaged. SAVO performance data from 26 trials were averaged resulting in one total score of deducted-points, reaction time, and decision time. Finally, the average positive and negative affect values were calculated for the pre-stress, the stress and post-stress periods at *T1* and *T2*. In all analyses mean differences between the pre-stress and the stress, as well as between the post-stress and stress phases respectively, were calculated and entered in repeated analyses of variance (ANOVA). When a significant interaction effect was found two post hoc two-tailed *t*-tests were conducted separately for the two groups comparing means for the dependent variable at T1 and T2. Consequently, alphas for these post hoc tests were adjusted to $\alpha = 0.025$. Since interaction effects were identified before by the ANOVA procedure, we used the *t*-test only to locate mean differences.

3. Results

3.1. Demographic variables

Table 2 displays means and standard deviations for demographic and socioeconomic background variables for both experimental groups. Age differed significantly between groups [$t(32) = 2.63$, $p = 0.02$, $d = 0.08$]. Natural science children ($M = 8.00$ years, $SD = 0.65$) were older than music group children ($M = 7.43$ years, $SD = 0.76$). Nevertheless, groups did not differ with regards to their educational level [$t(32) = 0.56$, $p = 0.58$], parental income [$t(21) = 0.68$, $p = 0.51$], cultural practice [$t(25) = 0.69$, $p = 0.50$], parents' educational achievements [father: $t(21) = 1.31$, $p = 0.21$; mother: $t(23) = 1.74$, $p = 0.10$], marks for social behavior [$t(26) = 0.81$, $p = 0.42$], or migration background (not displayed in the table; $\chi^2 = 1.40$; $p = 0.24$; mothers: $\chi^2 = 1.59$; $p = 0.45$; fathers: $\chi^2 = 4.01$; $p = 0.14$). Thus, both groups did not differ with respect to any variable but age. However, due to this group difference in age we repeated all ANOVAs with age as a covariate in order to examine in how far the effects might have been affected by the variance in age.

Table 2
Means (and SD) of the Musical Training Group (MT) and the Natural Science Training group (NST) for the demographic and socioeconomic variables at T1.

	MT	NST			
	M (SD)	*M (SD)*	*df*	*t*-value	*p*
Age	7.43 (0.76)	8.00 (0.65)	32	2.63	0.02
Education level (years)	2.50 (0.52)	2.60 (5.0)	32	0.56	0.58
Social behavior (grades)	1.71 (0.58)	2.18 (0.78)	26	0.81	0.42
Cultural practice	3.00 (1.10)	2.69 (1.20)	25	0.69	0.50
Parental income	1.90 (1.73)	1.46 (1.39)	21	0.68	0.51
Educational achievement (f)	5.00 (1.93)	3.93 (1.83)	21	1.31	0.21
Educational achievement (m)	4.20 (1.55)	2.87 (2.07)	23	1.74	0.10

Table 3
Means (and SD) of the SAVO, HR, SBP, DBP, PANAS and Cortisol measures for music training (MT) and natural science training (NST) groups for baseline, treatment and post treatment phase at both time points (T1-T2).

Measures	Music (n = 14)				Natural sciences (n = 20)			
	T1		T2		T1		T2	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
SBP pre-stress	101.50	(8.75)	107.31	(11.20)	105.88	(12.39)	111.32	(7.20)
SBP stress	107.31	(11.20)	111.38	(8.19)	106.11	(15.57)	108.72	(7.40)
SBP post-stress	100.88	(7.67)	108.10	(10.81)	100.05	(7.46)	108.23	(7.94)
DBP pre-stress	60.76	(8.20)	67.21	(7.28)	65.46	(7.50)	70.56	(6.57)
DBP stress	66.70	(6.74)	70.65	(7.39)	66.74	(9.81)	70.77	(6.68)
DBP post-stress	59.23	(8.30)	68.17	(10.60)	62.09	(6.71)	69.73	(6.37)
HR pre-stress	92.67	(11.23)	85.67	(9.82)	91.45	(10.53)	83.48	(9.57)
HR stress	93.29	(8.83)	88.63	(8.32)	90.23	(9.85)	85.30	(10.09)
HR post-stress	92.70	(11.17)	85.89	(10.42)	89.71	(8.60)	83.76	(8.95)
Cortisol pre-stress	5.08	(2.29)	6.42	(3.05)	6.23	(2.97)	8.15	(5.47)
Cortisol stress	5.95	(3.78)	8.09	(5.02)	8.01	(4.65)	7.33	(3.43)
Cortisol post-stress	5.04	(3.12)	7.41	(5.33)	7.44	(4.65)	6.43	(2.58)
PA pre-stress	1.41	(0.45)	1.19	(0.28)	1.25	(0.32)	1.28	(0.36)
PA stress	1.56	(0.33)	1.19	(0.46)	1.21	(0.40)	1.24	(0.35)
PA post-stress	1.57	(0.28)	1.21	(0.43)	1.28	(0.53)	1.16	(0.48)
NA pre-stress	0.20	(0.21)	0.07	(0.10)	0.18	(0.16)	0.13	(0.25)
NA stress	0.21	(0.30)	0.17	(0.21)	0.15	(0.17)	0.20	(0.24)
NA post-stress	0.17	(0.23)	0.11	(0.22)	0.12	(0.16)	0.07	(0.12)
SAVO RT (ms)	391	(52)	366	(65)	382	(138)	309	(35)
SAVO DT (ms)	9038	(4739)	6977	(2594)	6917	(2594)	5718	(2017)
SAVO SP	61.16	(28.21)	54.82	(26.71)	43.15	(25.26)	61.94	(29.02)

3.2. Responses to the task

To verify whether the SAVO task was able to elicit a physiological stress response paired *t*-tests (one-sided) for the cardiovascular (SBP, DBP and HR) and neurohumoral (salivary cortisol) measures at baseline (T1) were conducted. Results revealed significant differences between pre-stress and stress periods for DBP [$t(33) = 1.95$, $p < 0.05$, $d = 0.71$] and salivary cortisol [$t(30) = 1.87$, $p < 0.05$ $d = 0.54$]. Moreover, comparing means between stress and post-stress periods showed significant differences for all measures but HR [SBP: $t(33) = 2.78$, $p = 0.01$, $d = 0.97$; DBP: $t(33) = 4.32$, $p < 0.001$, $d = 1.50$; salivary cortisol: $t(31) = 1.81$, $p < 0.05$, $d = 0.67$; HR: $t(33) = -0.53$, $p = 0.30$]. Since the level of physiological arousal was initially higher than after children had accomplished the task we refer the arousal during the task to the post task measures as the more valid operationalization of a stress response (see Table 3). The heightened pre-task level probably was caused by the unfamiliarity of the situation and the excitement of the children about what they expected to come.

Significantly greater negative affects were reported for the stress-period compared to pre-stress [$t(33) = 2.80$, $p = 0.008$, $d = 0.98$] and post-stress measures [$t(33) = 2.08$, $p < 0.05$, $d = 0.72$] at T2 but not at T1, indicating that children showed an increase in subjective stress during the second time point only. According to the reported positive affects (PANAS), no differences between pre-stress, stress and post-stress measures were found at both time points. Taken together, these results suggest that the SAVO task induced a physiological stress response across all groups at both time points but subjectively the stress was experienced only at T2 when the children were older and more familiar with the task.

3.3. Development of task responses

Data were further analysed using repeated measure Analyses of Variance (ANOVA), which were performed on two dependent time points in this study. The quasi-experimental design was mixed model, with Group (music vs. natural science training) as the between-subjects factor and Time (T1; T2) as the within-subjects factor. Preconditions for conducting ANOVAs were tested (normality, Box's M Test of equality of covariates and Levene's Test of Equality of Variance) and were met in all cases.

No significant main effects for group or interaction effects were found for the reaction time (RT, all $Fs < 2.20$, $p > 0.15$) and for the decision time measure (DT, all $Fs < 4.32$, $p > 0.05$) of the SAVO task. This indicates that both groups were similarly engaged in the task and showed no differences in how strong children were decided to subtract the chosen amount of point from their opponents account. However, there was a main effect of time for the RT [$F(1, 29) = 5.57$, $p = 0.03$, $\eta_p^2 = 0.16$] and the DT [$F(1, 29) = 4.61$, $p = 0.04$, $\eta_p^2 = 0.14$] measures showing that children of both groups took less time to respond and to decide for the SAVO tasks at T2. These findings indicate a training effect on the SAVO tasks from T1 to T2.

With respect to the provocation of aggressive behavior measures no main effects (all $Fs < 1.5$, $p > 0.23$) but a significant interaction was found [$F(1, 29) = 5.69$, $p = 0.024$, $\eta_p^2 = 0.16$]. Using multiple *t*-tests for subsequent comparisons of means indicated natural science training children showed a significant increase in their aggression behavior scores from T1 to T2 [$t(16) = 2.54$, $p < 0.05$, $d = 1.27$], whereas the music children group did not

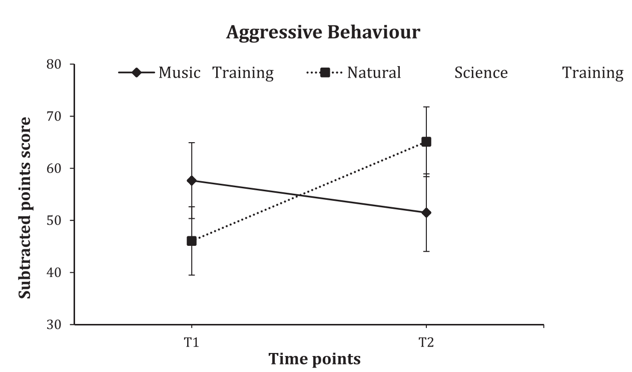

Fig. 1. Mean performances of aggressive behavior for the music training and the natural science training groups at baseline (T1) and after one and a half years of intervention (T2). Error falgs indicate Standard Errors of Means (SEM).

change significantly [$t(13) = 0.85$, $p > 0.05$]. This interaction remains statistically significant even if age was entered into the ANOVA as a covariate [$F(1, 28) = 4.51$, $p < 0.05$]. Fig. 1 depicts the disordinale Group × Time interaction for points subtracted by the children from their virtual opponents' account. Over the time course of one and a half year, children in the natural science group showed an increase in their aggressive behavior as operationalized by point reductions from their virtual opponents account while children in the music group did not. The latter rather showed a small decrease in their aggressive reaction to provocations.

The final set of analyses addressed the development of emotional (positive and negative affect ratings), cardiovascular (SBP, DBP and HR) and neurohumoral (salivary cortisol) stress responses to the SAVO task during the time course of 18 months. ANOVAs with response values (stress measures minus post-stress measures) as dependent variables and group (music vs. natural science) as between subjects factor and Time (T1, T2) as repeated measure revealed no main effects or interactions for SBP (all $Fs < 2.63$, $p > 0.12$), HR (all $Fs < 1.66$, $p > 0.21$), and salivary cortisol measures (all $Fs < 0.19$, $p > 0.65$). For DBP only a main effect of time was found indicating that independent of the group all children showed higher DBP responses at T1 compared to T2 [$F(1, 32) = 6.07$; $p = 0.02$; $\eta_p^2 = 0.16$]. However, this effect disappears if age is included as a covariate into the analysis [$F(1, 31) = 0.72$; $p = 0.40$], indicating that the appearance of this effect might be due to a confound of age and group. No other effects where found for DBP (all $Fs < 1.4$, $p > 0.24$).

Furthermore, for positive (all $Fs < 1.2$, $p > 0.29$) and negative (all $Fs < 1.41$, $p > 0.24$) affect ratings no main effects or interactions were found either.

4. Discussion and conclusions

This study investigated psychophysiological and behavioral responses to provocation of reactive aggression in primary school children receiving extended music training for 18 months. We found limited evidence for our hypotheses that music education would diminish stress responses.

In particular, our results indicate the successful induction of stress in both groups at both stress sessions spaced 18 months apart. Children of both the treatment and the control group showed increased blood pressures as well as heightened salivary concentrations of cortisol in response to the SAVO task. These findings confirm the validity of the SAVO task as a stressor and extend previous findings showing an association between anger and cardiovascular measures (Bongard & Al'Absi, 2003; 2005).

Concerning our first hypothesis, children receiving extended music or natural science training showed different patterns of reactive aggressive behaviors. Over the time course of 18 months children in the music group showed no significant change in this measure whereas children in the control group showed an increase. One interpretation of these findings is that children in the music group were stable in their coping strategies whereas the natural science children showed reduced capacities to withstand the provocation of aggressive behavior. These findings are reminiscent of a longitudinal study with older children (Lindblad et al., 2007). These authors observed that in the absence of changes of psychological variables at group level, music children showed lower cortisol concentrations after one year. Although the present study revealed a significant interaction (group × time), which appeared to favour musically trained children, this observation must be treated with caution due to the disordinal nature of this interaction (Shaffer, 1991).

Furthermore, children in the music and in the science group differed with respect to age but analyses of covariance indicated that the found effects remain even if the common variance of age and the dependant measures was eliminated.

There were no significant long-term effects neither of group nor of time of measurement on any of the physiological responses (HR, SBP, DBP, Cortisol), nor self-reported positive and negative affects. However, children in both groups showed similar significant increases in SBP and DBP measures and self-reported negative affects during the SAVO task. Contrary to our second hypothesis, music training had no beneficial effect on physiological responses to the provocation task.

Taken together, these findings shed new light on the psychophysiological long-term effects of specialized education programs in naturalistic instructional settings. While most studies in the past have focused on positive behavioral variables such as intelligence (e.g., Schellenberg, 2004), and cognitive competences in various domains (e.g., Roden et al., 2012, Roden, Könen, et al., 2014, Roden, Grube, et al., 2014), to our knowledge, this is the first study to look at the development of more less favoured aspects of emotional reactions over a prolonged period of time. However, although the duration of the intervention of 1.5 years exceeds previous studies, it might still be too short to elicit the expected changes in psychophysiological responses. Future studies might consider even longer periods of interventions in order to assess the socio-emotional and physiological impact of music training.

Changes in the physiological measures before and during the SAVO task were found to be of a rather small amount. For example, Bongard and Al'Absi (2003) and Chafin, Roy, Gerin, and Christenfeld (2004) reported average changes in the HR during a stress test (mental arithmetic task) of 10.4 bpm and 19.5 mmHg for the SBP or 13.4 mmHg for the DBP. Kirschbaum, Bartussek, and Strasburger (1992) showed an increase of cortisol levels by 6–9 nmol/l during a social stress test for adults. In a follow-up study, Buske-Kirschbaum et al., (1997) confirmed these results with regard to the changes of cortisol levels for children. Hence the SAVO task might be less efficient in the provocation of stress if compared to the results reported in the literature. Further research might focus on more challenging treatments to induce greater stress responses to find out more about the effects of music training on physiological stress responses in primary school children.

Finally, salivary cortisol measures at pre-task, task and post-task phases were limited to a time frame of one school lesson (45 min). Thus, the salivary cortisol level at pre-task was measured 5 min before the provocation of aggressive behavior started. Ten minutes after the beginning of the SAVO treatment test, the task measure of salivary cortisol level was taken, followed by a post-task cortisol measure after 20 min after the SAVO stress period ended. Regardless of the fact that the design of the study takes into account that the cortisol level reaches its maximum between 10 and 30 min after stress induction (Khalfa, Bella, Roy, Peretz & Lupien, 2003), a longer interval between the task and the post-task measurement of cortisol levels might lead to stronger increases or decreases of cortisol levels in both groups. Moreover, previous research focused on the change of cortisol level by measuring its responses directly before and after group singing (Beck, Cesario, Yousefi, & Enamoto, 2000; Fancourt et al., 2016; Kreutz, Bongard, Rohrmann, Hodapp, & Grebe, 2004). Thus measuring salivary cortisol directly before, during and after instrumental playing – while aggressive behavior was provoked – might lead to stronger changes in the cortisol levels as well. Further research will be necessary to clarify these assumptions.

4.1. Limitations

This study is limited in several ways. Firstly, the study is based on a quasi-experimental design observing groups in their natural environment at school. Although we evaluated demographic and socioeconomic variables such as educational level, social behavior, parental income and educational achievement, to determine any systematic bias prior to the intervention, interpretations must be made with caution. Future research could employ experimental longitudinal methods to enhance our findings. In the present study the natural science group served as control intervention to contrast the effects of extended music education. However, both music and science education groups may also differ in other aspects than just playing a musical instrument or not. For example, learning to play a musical instrument might involve more physical activity than natural science training or making music might be experienced more creative compared to practicing experimental procedures. Therefore, it would be beneficial to consider further alternative educational programs like sports or drama as control conditions for effects of musical education in future research. Secondly, the small sample size does not allow the examination of the impact of specific kinds of aggressive behavior. Larger scale studies would aid in the generalizability of these findings. However, a priori power analyses showed that the sample size of the current study was large enough for the expected effects. Thirdly, children in the treatment group were only involved in a school-based music program. Therefore, generalization to other music intervention programs – especially to those offering individual, private music lessons – has to be treated with caution. Fourthly, we only found evidence for a relative differentiation of reactive aggressive behavior between the groups in the behavioral measures of the provoked reactive aggressive behavior but not in the physiological stress responses. Given the fact that all our measures were taken during school time between 8 a.m. and 12 p.m., one might argue that the observed salivary cortisol levels across the two groups were affected by its diurnal rhythm (Kirschbaum & Hellhammer, 1994). However, there were no differences in time for the onset of the cortisol measurements at both time points. Furthermore, the decreases in salivary cortisol levels as observed for both groups after the SAVO task – together with the decrease of SBP, DBP and negative affects after the SAVO task – made the diurnal effects of cortisol seem less substantial for the perceived changes in the current study. Moreover we usually see no stronger decreases for salivary cortisol levels from stress to post-stress than we see increases from pre-stress to stress except for the natural science group at T2, which also makes it less likely that diurnal effects were of great influence in this study. Nevertheless, further studies are necessary to clarify the impact of the SAVO task on the salivary cortisol level of school-aged children. Finally, one might argue that the increased aggressive behavior of children in the natural science group after 18 months depends on different conceptions of learning and teaching between both intervention groups (e.g. López-Íñiguez & Pozo, 2014). In the current study, music and natural science training was organized in small groups, in which the teaching and learning process focused on students' collaborative learning and social communication. Hence, systematic influences due to different learning and teaching conceptions seem less likely to explain the observed results.

5. Conclusion

The present study provides preliminary evidence that musically trained children demonstrate a lower increase in reactive aggressive behavior when compared to the control group, and this even when the physiological stress responses increased across groups. Considering the explorative nature of the study, we conclude that school-based music training can be beneficial in the sense of preventing increases in aggressive behavior, thus supporting previous investigations. However, more research is necessary to ascertain positive effects of music training on well-being and health in primary school children.

Acknowledgments

This study is supported by a grant from the German Federal Ministry of Education and Research, which was awarded to the fifth and last authors (No. 01KJ0808). The founding source had no involvement on the study design, the collection, analysis and interpretation of data, the writing of the report or the decision to submit the article for publication. We like to thank all our student assistance involved in the study for their support during data acquisition and processing. Particularly we like to thank all children, parents, teachers and administrators for their support.

Disclosure statement/Declaration of interests

The second author (FDZ) was the recipient of an unrestricted award donated by the American Psychiatric Association (APA), the American Psychiatric Institute for Research and Education (APIRE) and AstraZeneca (Young Minds in Psychiatry Award). He has also received a travel stipend donated by the GlaxoSmithKline Foundation. He has also received an unrestricted educational grant, travel support and speaker honoraria by Shire Pharmaceuticals, Germany. In addition, he has received editorial fees from Co-Action Publishing (Sweden). FDZ does not report a conflict of interest with regards to the present publication.

References

Achenbach, T. (1991). *Child behavior checklist/4–18*. Burlington: University of Vermont.

Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of Experimental Child Psychology, 83*(2), 111–130.

Beck, R., Cesario, T., Yousefi, A., & Enamoto, H. (2000). Choral singing, performance perception, and immune system changes in salivary immunoglobulin A and cortisol. *Music Perception, 18*(1), 87–106.

Bongard, S., & Al'Absi, M. (2003). Domain-specific anger expression assessment and blood pressure during rest and acute stress. *Personality and Individual Differences, 34*(8), 1383–1402.

Bongard, S., & Al'Absi, M. (2005). Domain-specific anger expression and blood pressure in an occupational setting. *Journal of Psychosomatic Research, 58*(1), 43–49.

Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Valtin, R., Voss, A., et al. (2005). *IGLU: Skalenhandbuch zur Dokumentation der Erhebungsinstrumente [Progress in international reading literacy study]*. Münster: Waxmann.

Burnett, P. C. (1996). Gender and grade differences in elementary school children's descriptive and evaluative self-statements and self-esteem. *School Psychology International, 17*, 159–170.

Buske-Kirschbaum, A., Jobst, S., Wustmans, A., Kirschbaum, C., Rauh, W., & Hellhammer, D. (1997). Attenuated free cortisol response to psychosocial stress in children with atopic dermatitis. *Psychosomatic Medicine, 59*(4), 419–426.

Chafin, S., Roy, M., Gerin, W., & Christenfeld, N. (2004). Music can facilitate blood pressure recovery from stress. *British Journal of Health Psychology, 9*(3), 393–403.

Choi, A. N., Lee, M. S., & Lee, J. S. (2010). Group music intervention reduces aggression and improves self-esteem in children with highly aggressive behavior: A pilot controlled trial. *Evidence-Based Complementary and Alternative Medicine, 7*(2), 213–217. http://dx.doi.org/10.1093/Ecam/Nem182.

Costa-Giomi, E. (2004). Effects of three years of piano instruction on children's academic achievement, school performance and self-esteem. *Psychology of Music, 32*(2), 139–152.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers, 28*(1), 1–11.

Fancourt, D., Williamon, A., Carvalho, L. A., Steptoe, A., Dow, R., & Lewis, I. (2016). Singing modulates mood, stress, cortisol, cytokine and neuropeptide activity in cancer patients and carers. *ecancermedicalscience, 10*.

Frankenberg, E., Fries, K., Friedrich, E. K., Roden, I., Kreutz, G., & Bongard, S. (2014). The influence of musical training on acculturation processes in migrant children. *Psychology of Music*. http://dx.doi.org/10.1177/0305735614557990.

Hallam, S. (2010). The power of music: Its impact on the intellectual, social and personal development of children and young people. *International Journal of Music Education, 28*(3), 269–289.

Ho, Y. C., Cheung, M. C., & Chan, A. S. (2003). Music training improves verbal but not visual memory: Cross-sectional and longitudinal explorations in children. *Neuropsychology, 17*(3), 439.

Ho, P., Tsao, J. C., Bloch, L., & Zeltzer, L. K. (2011). The impact of group drumming on social-emotional behavior in low-income children. *Evidence-Based Complementary and Alternative Medicine, 1*, 1–14. doi: 25070810.1093/ecam/neq072.

Khalfa, S., Bella, S. D., Roy, M., Peretz, I., & Lupien, S. J. (2003). Effects of relaxing music on salivary cortisol level after psychological stress. *Annals of the New York Academy of Sciences, 999*(1), 374–376.

Kirschbaum, C., Bartussek, D., & Strasburger, C. (1992). Cortisol responses to psychological stress and correlations with personality traits. *Personality and Individual Differences, 13*(12), 1353–1357.

Kirschbaum, C., & Hellhammer, D. H. (1994). Salivary cortisol in psychoneuroendocrine research: Recent developments and applications. *Psychoneuroendocrinology, 19*(4), 313–333.

Knox Anderson, H., & Rickard, N. S. (2007). A preliminary examination of short-term effects of an active class-based music program on young adolescents' self-esteem and anger expression. *Australian Journal of Music Education, 1*, 4–16.

Koger, S. M., Chapin, K., & Brotons, M. (1999). Is music therapy an effective intervention for dementia? A meta-analytic review of literature. *Journal of Music Therapy, 36*(1), 2–15.

Kreutz, G., Bongard, S., Rohrmann, S., Hodapp, V., & Grebe, D. (2004). Effects of choir singing or listening on secretory immunoglobulin A, cortisol, and emotional state. *Journal of Behavioral Medicine, 27*(6), 623–635.

Krohne, H. W., Egloff, B., Kohlmann, C.-W., & Tausch, A. (1996). Untersuchungen mit einer deutschen Version der "Positive and Negative Affect Schedule"(PANAS). *Diagnostica, 42*, 139–156.

Laohawattanakun, J., Chearskul, S., Dumrongphol, H., Jutapakdeegul, N., Yensukjai, J., Khumphan, N., … Thangnipon, W. (2011). Influence of music training on academic examination-induced stress in Thai adolescents. *Neuroscience Letters, 487*(3), 310–312.

Lee, Y.-S., Lu, M.-J., & Ko, H.-P. (2007). Effects of skill training on working memory capacity. *Learning and Instruction, 17*(3), 336–344. http://dx.doi.org/10.1016/j.learninstruc.2007.02.010.

Lindblad, F., Hogmark, A., & Theorell, T. (2007). Music intervention for 5th and 6th graders - Effects on development and cortisol secretion. *Stress and Health, 23*(1), 9–14. http://dx.doi.org/10.1002/Smi.1109.

Linnemann, A., Ditzen, B., Strahler, J., Doerr, J. M., & Nater, U. M. (2015). Music listening as a means of stress reduction in daily life. *Psychoneuroendocrinology, 60*, 82–90.

López-Íñiguez, G., & Pozo, J. I. (2014). The influence of teachers' conceptions on their students' learning: Children's understanding of sheet music. *British Journal of Educational Psychology, 84*(2), 311–328.

Ostermeier, C., Prenzel, M., & Duit, R. (2010). Improving science and mathematics instruction - The SINUS-project as an example for reform as teacher professional development. *International Journal of Science Education, 32*(3), 303–327.

Pelham, W. E., Milich, R., Cummings, E. M., Murphy, D. A., Schaughency, E. A., & Greiner, A. R. (1991). Effects of background anger, provocation, and methylphenidate on emotional arousal and aggressive responding in attention-deficit hyperactivity disordered boys with and without concurrent aggressiveness. *Journal of Abnormal Child Psychology, 19*(4), 407–426.

Rickard, N. S., Appelman, P., James, R., Murphy, F., Gill, A., & Bambrick, C. (2013). Orchestrating life skills: The effect of increased school-based music classes on children's social competence and self-esteem. *International Journal of Music Education, 31*(3), 292–309.

Rickard, N. S., Bambrick, C. J., & Gill, A. (2012). Absence of widespread psychosocial and cognitive effects of school-based music instruction in 10–13-year-old students. *International Journal of Music Education, 30*(1), 57–78.

Roden, I., Grube, D., Bongard, S., & Kreutz, G. (2014). Does music training enhance working memory performance? Findings from a quasi-experimental longitudinal study. *Psychology of Music, 42*(2), 284–298. http://dx.doi.org/10.1177/0305735612471239.

Roden, I., Könen, T., Bongard, S., Frankenberg, E., Friedrich, E. K., & Kreutz, G. (2014). Effects of music training on attention, processing speed and cognitive music abilities—Findings from a longitudinal study. *Applied Cognitive Psychology, 28*(4), 545–557.

Roden, I., Kreutz, G., & Bongard, S. (2012). Effects of a school-based instrumental music program on verbal and visual memory in primary school children: A longitudinal study. *Frontiers in Psychology, 3*(572), 1–9.

Sandstrom, G. M., & Russo, F. A. (2010). Music hath charms the effects of valence and arousal on recovery following an acute stressor. *Music and Medicine, 2*(3), 137–143.

Savan, A. (1999). The effect of background music on learning. *Psychology of Music, 27*(2), 138–146.

Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychological Science, 15*(8), 511–514.

Shaffer, J. P. (1991). Probability of directional errors with disordinal (qualitative) interaction. *Psychometrika, 56*(1), 29–38.

Southgate, D. E., & Roscigno, V. J. (2009). The impact of music on childhood and adolescent achievement. *Social Science Quarterly, 90*(1), 4–21.

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding speech prosody: Do music lessons help? *Emotion, 4*(1), 46.

Varela, W., Abrami, P. C., & Upitis, R. (2014). Self-regulation and music learning: A systematic review. *Psychology of Music*, 0305735614554639.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063.

Yehuda, N. (2011). Music and stress. *Journal of Adult Development, 18*(2), 85–94.

Zepf, F. D., Stadler, C., Demisch, L., Schmitt, M., Landgraf, M., & Poustka, F. (2008). Serotonergic functioning and trait-impulsivity in attention-deficit/hyperactivity-disordered boys (ADHD): Influence of rapid tryptophan depletion. *Human Psychopharmacology: Clinical and Experimental, 23*(1), 43–51.